



UNIVERSITY
OF MANITOBA

Off-line LC-MALDI techniques for the identification of proteins, post-
translational modifications and protein-protein interactions in proteomic
studies

by

Vincent C. Chen

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements of the degree of

Doctorate (PhD) of Science

Department of Chemistry

University of Manitoba

Winnipeg

Copyright © by Vincent C. Chen

THE UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

Off-line LC-MALDI techniques for the identification of proteins, post-translational modification and protein-protein interaction in proteomic studies

BY

Vincent C. Chen

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of
Manitoba in partial fulfillment of the requirement of the degree**

Doctorate (PhD) of Science

Vincent C. Chen © 2008

Permission has been granted to the University of Manitoba Libraries to lend a copy of this thesis/practicum, to Library and Archives Canada (LAC) to lend a copy of this thesis/practicum, and to LAC's agent (UMI/ProQuest) to microfilm, sell copies and to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

Abstract

Cells process an array of intra- and extra-cellular signals needed to coordinate a variety of activities such as cell division, movement and differentiation. However, upon closer inspection of these networks, signal transduction can be broken down to a series of protein-protein interactions which are often modulated by the presence (or absence) of sequence specific modifications. While deciphering these events is extremely important, proteome-level analyses of these networks has proven to be a challenge due to the complex nature of biological samples. Although large-scale automation has revolutionized the field of proteomics, little ground has been gained in the development of robust microvolume approaches for one of the most commonly performed tasks within the life sciences, the purification of protein samples. To address these problems, we have set out to develop novel separation techniques for the characterization of protein-protein interactions and protein phosphorylation by mass spectrometry. To aid studies, a vacuum-driven LC device was developed for the separation and MALDI deposition of subpicomole amounts of material for MS and MS/MS analyses. To demonstrate utility of this device, we set out to identify novel protein-protein interactions of zonula occludens-1 (ZO-1), a member of the MAGUK family of membrane-associated adaptor proteins. Using a GST-fusion protein incorporating the PDZ1 domain of ZO-1, interactions involving the protein domain were pulled-down for proteomic identification. Over the course of the study, several cortical scaffolding proteins were identified, including alpha-

actinin-4. Further in-depth analysis of the interaction between ZO-1 and alpha-actinin-4 demonstrated the interaction existed within a wide variety of cell and tissue types. Interestingly, alpha-actinin-1, a protein having high sequence identity (86%) to alpha-actinin-4, demonstrated no affiliation with the ZO-1 protein. In the final chapter, an analytical method to increase phosphopeptide detection was developed and demonstrated based upon the off-line LC-MALDI platform. Here, a strategy employing LC retention time prediction was used to enhance the detection of low abundance phosphopeptides in a hypothesis-driven manner. While many aspects of the research presented are proof-of-concept in nature, developments of these methods are expected to help researchers uncover dynamic events associated with cell signal transduction.

Acknowledgements

I would first like to thank committee members Dr. Harry Duckworth and Dr. Joe O'Neil for providing guidance, input and critical evaluations over the years. I would also like to thank Dr. Michael Siu (York University) for serving as my external reviewer, and collaborators at the National Core Facilities for Proteomics, specifically, Dr. K.-H. Khoo, Dr. C.-C. Chou and Irene Hsieh. Many thanks to current and former members of the University of Manitoba, including Shaheen Shojania, Sergei Snovida, Julian Saba, Colin Lee, Victor Spicer, Keding Cheng, Yuwei Qian, Erika Lattova and faculty members Dr. Oleg Krokhin, Dr. Werner Ens, Dr. Kenneth Standing, Dr. Jim Charlton, and Dr. H.D. Gesser. I also express much gratitude to Dr. Xinbo Li, my co-supervisor, Dr. James Nagy, and my primary supervisor, Dr. H el ene Perreault, for introducing me to the areas of mass spectrometry and cell biology. Finally, I am most grateful for my family, my mother and father, and especially my wife Beverley and daughter Jaelyn, for their unconditional love and support.

Many thanks!

-V.Chen

For Beverley and Jaelyn

Abstract	iii
Acknowledgments	iv
Dedication	v
Table of Contents	vi
List of figures	xiii
List of tables	xviii
List of abbreviations	xx
Publications	xxv
1.0 General Introduction	1
1.1 Transcriptomics.....	3
1.2 Proteomics.....	5
1.2.1 Protein Identification.....	6
1.2.2 Protein Quantification.....	7
1.2.3 Protein-protein Interaction and Functional Proteomics.....	10
1.2.3.1 PDZ Domains.....	12
1.2.3.2 Functional Proteomic Approaches.....	16
1.2.4 Post-translational Modifications (PTM).....	20
1.3 General Methodology.....	25
1.3.1 Gel Electrophoresis.....	27
1.3.2 Liquid Chromatography.....	29
1.3.3 Principles of Mass Spectrometry and Instrumentation.....	32
1.3.3.1 Matrix-assisted Laser Desorption/Ionization.....	35

1.3.3.2 Electrospray Ionization.....	37
1.3.3.3 Quadrupole.....	39
1.3.3.4 Time-of-flight.....	40
1.3.3.4.1 Time-lag Focusing.....	42
1.3.3.4.2 Reflectron Ion Mirror.....	45
1.3.4.5 MALDI Quadrupole-quadrupole Time-of-flight.....	46
1.3.4.6 MALDI Time-of-flight/Time-of-flight.....	48
1.3.4.7 Protein identification by Peptide-mass Fingerprint (PMF)...	51
1.3.4.8 Protein identification by Tandem-MS.....	53
1.4 Rational for studies.....	59
1.5 References.....	68

2.0 Development of an off-line vacuum-LC/deposition device for

MALDI-MS and MS/MS.....	82
Author's Contributions to published materials.....	83
2.1 Abstract.....	85
2.2 Introduction.....	86
2.3 Experimental	
2.3.1 Materials.....	90
2.3.2 Peptides Used for Ion Suppression Study.....	91
2.3.3 Digestion of β -casein.....	91
2.3.4 Digestion of <i>E. coli</i> citrate synthase.....	91
2.3.5 SepDep Device.....	92

2.3.6 Hydrophobic MALDI Target Preparation.....	94
2.3.7 Matrix Preparation, Tilt-and-tap pre-spotting	95
2.3.8 Reverse Phase Microcolumns.....	95
2.3.9 Sample Loading and Separation.....	96
2.3.10 MALDI mass spectrometry.....	97
2.3.11 MALDI-MS/MS.....	97
2.4 Results and Discussion	
2.4.1 Design and Performance of the SepDep Device.....	99
2.4.2 Polybutadiene Hydrophobic Sample Barrier.....	100
2.4.3 Reduction of Ion Suppression Effects.....	102
2.4.4 Separation of Protein Digests.....	104
2.5 Conclusions.....	100
2.6 Acknowledgements.....	111
2.7 References.....	112
3.0 Development of Proteomic Approaches for the Identification and Characterization of PDZ-protein interactions.....	116
Author's contributions to published materials.....	117
3.1 Abstract.....	121
3.2 Introduction.....	122
3.3 Experimental	
3.3.1 Antibodies and Animals.....	125
3.3.2 Cell culture and tissue preparation.....	125

3.3.3 GST-PDZ domain fusion proteins.....	126
3.3.4 Pull-down of proteins for MS analysis.....	127
3.3.5 Mass Spectrometry.....	128
3.3.6 Pull-down for Confirmation of PDZ/ α -Actinin-4 Interaction....	130
3.3.7 Immunoprecipitation.....	131
3.3.8 Western-Blotting.....	132
3.3.9 Light microscope Immunofluorescence.....	132
3.4 Results	
3.4.0 Characterization of recombinant fusion proteins by MALDI-MS.....	134
3.4.1 MS Identification of α -Actinin-4 pull-down with PDZ1 Domain of ZO-1.....	136
3.4.2 Immunoblotting of α -Actinin-4 from Cultured Cells and Mouse Tissue.....	141
3.4.3 Co-IP of α -Actinin-4 with ZO-1.....	141
3.4.4 α -Actinin-4 Binding to the PDZ1 of ZO-1.....	144
3.4.5 Lack of α -Actinin1 Co-IP with ZO-1.....	147
3.4.6 Immunofluorescence Co-localization of α -Actinin-4 with ZO-1.....	147
3.5 Discussion and Conclusions.....	153
3.5.1 MS-based Identification of Protein Interactions.....	154
3.5.2 ZO-1/ α -Actinin-4 Interaction.....	156
3.5.3 Functional Implications.....	157

3.6 References.....	160
---------------------	-----

4.0 Development of Off-line HPLC MALDI-MS/MS Methods for High-

Content Phosphoproteomics using LC Retention Time

Prediction.....	166
Authors Contributions.....	167
4.1 Abstract.....	169
4.2 Introduction.....	170
4.3 Experimental	
4.3.1 Materials.....	174
4.3.2 Off-line High-performance Liquid Chromatography and MALDI Fraction Collection.....	174
4.3.3 Mass Spectrometry.....	176
4.3.4 Workflow and Data-dependent Analysis.....	177
4.3.5 Seven-protein Tryptic Digest Mixture.....	182
4.3.6 Myelin Basic Protein (MBP).....	184
4.3.7 Anti-phosphotyrosine Immunoprecipitation.....	188
4.4 Results and Discussion.....	199
4.4.1 Seven-protein Digest Mixture.....	201
4.4.2 Myelin Basic Protein.....	201
4.4.3 Anti-pY Immunoprecipitates: Abelson-interacting protein.....	202
4.5 Acknowledgements.....	206
4.6 References.....	207

5.0 Conclusions and Future Work.....	209
5.1 Introduction.....	210
5.2 Development of an Off-line Vacuum-LC/Deposition device for MALDI-MS and MS/MS.....	211
5.3 Development of Proteomic Approaches for the Identification And Characterization of Protein-protein Interactions modulated by PDZ-domains.....	215
5.4 Development of Off-line HPLC MALDI-MS/MS Methods for High-content Phosphoproteomics using LC Retention Time Prediction.....	223
5.6 References.....	224
6.0 Appendix: Method and Apparatus for Depositing Samples on a Target Surface.....	227
6.1 Abstract.....	229
6.2 Field of the invention.....	230
6.3 Background of the invention.....	230
6.4 Materials and Methods.....	231
6.5 Summary of the invention.....	236
6.5.1 Brief description of the drawings.....	237
6.5.2 Detailed descriptions of the invention.....	244
6.5.2 Claims.....	255

7.0 Supplementary Information: Targeted identification of phosphorylated peptides by off-line HPLC MALDI-MS/MS using LC Retention Time prediction.....260

8.0 Supplementary information: General Protocols and Stock Solutions.....283

List of figures

Figures: General Introduction

1.1	Central dogma of molecular biology.....	2
1.2	Technique for the relative quantification of proteins by SILAC.....	9
1.3	PDZs and PDZ-domain containing protein ZO-1.....	13
1.4	Identification of protein-protein interactions with GST-fusion protein.....	17
1.5	Identification of protein-protein interactions by Coimmunoprecipitation.....	20
1.6	Protein phosphorylation.....	22
1.7	Commonly used techniques in proteomics.....	26
1.8	Commonly used MALDI matrices: DHB and α -Cyano.....	36
1.9	MALDI ionization process.....	36
1.10	Electrospray ionization coupled to HPLC.....	38
1.11	Quadrupole mass selector.....	39
1.12	Time-of-flight mass analyzer.....	41
1.13	Principles of time-lag focusing.....	44
1.14	DE Reflectron time-of-flight mass spectrometer.....	45
1.15	MALDI Quadrupole-quadrupole time-of-flight mass spectrometer (MALDI-QqTOF).....	48
1.16	MALDI Time-of-flight/time-of-flight mass spectrometer (MALDI TOF/TOF-MS).....	50
1.17	Principles of peptide-mass fingerprinting (PMF).....	52

1.18	Peptide fragmentation nomenclature	55
1.19	MS/MS data for the analysis of peptide sequences.....	56
1.20	Strategy for protein identification using MS/MS data.....	58
1.21	Data-dependent analysis by on-line LC-MS/MS.....	62
1.22	A typical off-line LC MALDI interface.....	65

Part 2 Figures: Development of an Off-line Vacuum-LC/Deposition device for MALDI-MS and MS/MS

2.1	Components of the SepDep device.....	93
2.2	Hydrophobic sample barrier for MALDI-MS.....	94
2.3	SepDep protocol.....	98
2.4	Reduction of MALDI ion suppression effects.....	103
2.5	Representative spectra from off-line vacuum chromatography of citrate synthase tryptic peptides.....	105
2.6	CS sequence coverage by SepDep MALDI-MS.....	106
2.7	Identification of β -casein tetra-phosphorylated peptide.....	108
2.8	MS and MS/MS of β -casein phosphorylated peptides.....	109
2.9	SepDep MS detection of β -casein peptide m/z 1981.9.....	110

Part 3 Figures: Development of proteomic approaches for the identification and characterization of protein-protein interactions modulated by PDZ-domains

3.0	Characterization of GST-fusion proteins by MS.....	135
-----	--	-----

3.1	MS identification of α -actinin-4 following pull-down with GST-fusion protein containing the PDZ1 domain of ZO-1.....	139
3.2	Immunoblots of cultured cells and mouse tissues probed with anti- α -actinin-4 antibodies.....	142
3.3	Co-IP of α -actinin-4 and ZO-1.....	143
3.4	Pull-down analyses showing α -actinin-4 interaction with the PDZ1 domain of ZO-1.....	146
3.5	Lack of α -actinin-1 association with ZO-1.....	148
3.6	Immunofluorescence labeling of ZO-1 and α -actinin-4 in cultured HeLa cells and C6 glioma cells.....	149
3.7	Double immunofluorescence labeling of ZO-1 and α -actinin-4 in tissues of adult mouse.....	152

**Part 4 Figures: Development of off-line HPLC MALDI-MS/MS methods
for high-content phosphoproteomics using LC retention time prediction**

4.1	HPLC-UV trace.....	176
4.2	Retention time calibration using myoglobin peptides.....	179
4.3	Search Peak interface.....	181
4.4	MS/MS of casein phosphopeptides from α - and β -casein within 7 protein digest mixture.....	183
4.5	MALDI-MS/MS spectra of tryptic phosphopeptides of myelin basic protein m/z 1571.8 (A) and 1419.7 (B).....	186
4.6	ABI phosphopeptide TPPVNNPPQVPSHpYAPNYPIGHPKR.....	192

4.7	LC retention time calibration for peptides obtained from S2 cell protein Immunoprecipitates.....	193
4.8	Off-line HPLC MALDI-MS Fraction 43.....	196
4.9	Off-line HPLC MALDI-MS Fraction 46.....	197
4.10	MS/MS spectra of ABI phosphopeptides.....	198
4.11	Relative deviations of experimentally confirmed phosphopeptides and retention times compared to predicted values.....	200

Part 5 Figures: Conclusions and Future Work

5.1	ZO-1/ α -actinin-4 mediated transcriptional regulation.....	220
5.2	Identification of protein-protein interactions of ZO-1 PDZ1, PDZ2, and PDZ3 domains by SILAC.....	222

**Part 6 Appendix Figures: Method and Apparatus for Depositing Samples
on a Target Surface**

6.1	Schematic representation of the SepDep II Vacuu-flow sample preparation module for off-line LC MALDI MS.....	232
6.2	Images of the SepDep II prototype.....	233
6.3	Identification of tetra-phosphorylated peptide.....	235

Device Embodiments:

Figure 6.4.....	239
Figure 6.5A.....	240

Figure 6.5B.....	241
Figure 6.7A, B, C.....	242
Figure 6.8.....	249
Figure 6.9A.....	251
Figure 6.9B.....	252

List of Tables

Part 1 Tables: General Introduction

1.1	Examples of PDZ domains classified by C-terminal ligands.....	15
-----	---	----

Part 3.0 Tables: Development of proteomic approaches for the identification and characterization of protein-protein interactions modulated by PDZ-domains

3.1	Identification of ZO-1-associated scaffolding network by MALDI-MS.....	137
3.2	Sequence alignment between non-muscle α -actinin-1 and α -actinin-4.....	140

Part 4.0 Tables: Development of off-line HPLC MALDI-MS/MS methods for high-content phosphoproteomics using LC retention time prediction

4.1	Calibration and LC retention time data for phosphorylated proteins α - and β -casein within the 7 protein digest mixture.....	178
4.2	LC retention data of α - and β -casein phosphopeptides.....	184
4.3	LC retention time variance for tryptic phosphopeptides from myelin basic protein.....	187
4.4	Top 5 proteins identified within anti-pY immunoprecipitates of S2 cell proteins by automated (intensity-driven) Data-dependent analysis.....	189

4.5	SSRcal calibration for Abelson-interacting protein.....	193
4.6	MS/MS inclusion list indicating fraction location, <i>m/z</i> of candidate Sequences of ABI.....	194

Part 8 Supplementary Information Tables: General Protocols and Stock Solutions

8.0	SDS-PAGE mini-gels.....	284
8.1a	SDS-PAGE separating gels.....	286
8.1b	SDS-PAGE stacking gel.....	286
8.2	Western-blot.....	287
8.3	In-gel digestion and peptide recovery for MALDI-MS.....	289
8.4	ZipTip™ and application to MALDI plate.....	290
8.5	Cell passage (HeLa, PC12).....	291
8.6	Immunohistochemistry – cultured cells.....	292
8.7	Immunohistochemistry – tissue	295
8.8	SDS-PAGE stock solutions.....	297
8.9	Immunohistochemistry stock solutions.....	299
8.10	Western-blot stock solutions.....	300

Abbreviations

ABI	Abelson-interacting protein
ACN	acetonitrile
α -cyano	α -cyano-4 hydroxycinnamic acid
ADH	alcohol dehydrogenase
Amu	atomic mass unit
Anti-pY	anti phosphotyrosine
ATP	adenosine triphosphate
C18, C8, C4	reverse phase chromatography
CAN	Canadian/Canada
CID	collision induced dissociation
CNS	central nervous system
Co-IP	co-immunoprecipitation
Cs	citrate synthase
Cx36	connexin36
Cx43	connexin43
Cx45	connexin45
Cy3	cyano3
°C	degree celsius
Da	Dalton
DE	delayed extraction
DHB	2,5-dihydroxybenzoic acid
2D-LC	2-dimensional liquid chromatography
DTT	dithiothreitol
2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
EDTA	Ethylamine diamine tetra acetate
ESDL-COOH	C-terminal sequence of α -actinin-1/-4
ESI	electrospray ionization

f	femto, x10 ⁻¹⁵
FA	formic acid
FITC	fluorescein isothiocyanate
GST	glutathione-s-transferase
h	hour
HCl	hydrochloric acid
H ₂ O	water
HPLC	high performance liquid chromatography
Hz	hertz, /second
IAA	iodoacetamide
ID	inner diameter
IgG	immunoglobulin G
IMAC	immobilized metal affinity chromatography
IP	immunoprecipitation
IPG	immobilized pH gradient
kDa	kiloDalton
kV	kilovolts
keV	kiloelectrovolts
L	liter
l	length
LC	liquid chromatography
μ	micro, x10 ⁻⁶
m	moles
M	molar
MALDI	matrix-assisted laser desorption/ionization

MAP	microtubule associated protein
MBP	myelin-basic protein
MgCl ₂	magnesium chloride
Mr	relative molecular Weight
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MQ	Milli-Q water
Mw	molecular weight
MWCO	molecular weight cut off
<i>m/z</i>	mass-to-charge ratio
n	nano, x10 ⁻⁹
NaCl	sodium chloride
NCBI	National Center for Biotechnology Information (USA)
nm	nanometer
nr	non-redundant
p	pico, x10 ⁻¹²
pS	phosphoserine
ps	picosecond
pT	phosphothreonine
pY	phosphotyrosine
PAGE	polyacrylamide gel-electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMF	peptide mass fingerprinting
pmol	picomol
PSD	post source decay
PTM	post-translational modification
PDZ	protein domains having homology with post-synaptic density, drosophila large disc and zonula occludens-1

Q	Quadrupole
q	quadrupole collision cell for MS/MS
QqTOF	quadrupole-quadrupole time-of-flight
RP	reverse phase
Sep-Dep	vacuum activated separation-deposition device
Sep-Dep II	second generation vacuum separation/deposition device
SCX	strong cation exchange
SDS	sodium dodecyl sulfate
SILAC	stable isotopic labelling in cell culture
SPE	solid phase extraction
S/N	signal-to-noise
<i>SSRCalc</i>	Sequence specific retention time calculator
TPCK	L-1-tosylamide-2-phenylethylchloromethyl ketone
TFA	trifluoroacetic acid
TIS	timed-ion-selector
TOF	time-of-flight
TOF/TOF	time-of-flight/time-of-flight, tandem time-of-flight
TSTw	buffer containing Tris-HCl, sodium chloride, Tween-20
.txt	text file format
X	any amino acid
ZO-1	zonula occludens-1

Abbreviations and masses of the 20 common amino acids*

Amino acid	Symbol		Monoisotopic (Da)	Average (Da)
Alanine	Ala	A	71.0371	71.0788
Arginine	Arg	R	156.1011	156.1876
Asparagine	Asn	N	114.0429	114.1039
Aspartic acid	Asp	D	115.0269	115.0886
Cysteine	Cys	C	103.0092	103.1448
Glutamine	Gln	Q	128.0586	128.1308
Glutamic acid	Glu	E	129.0426	129.1155
Glycine	Gly	G	57.0215	57.0520
Histidine	His	H	137.0589	137.1412
Isoleucine	Ile	I	113.0841	113.1595
Leucine	Leu	L	113.0841	113.1595
Lysine	Lys	K	128.0950	128.1742
Methionine	Met	M	131.0405	131.1986
Phenylalanine	Phe	F	147.0684	147.1766
Proline	Pro	P	97.0528	97.1167
Serine	Ser	S	87.0320	87.0782
Threonine	Thr	T	101.0477	101.1051
Tryptophan	Trp	W	186.0793	186.2133
Tyrosine	Tyr	Y	163.0633	163.1760
Valine	Val	V	99.0684	99.1326
Any Amino Acid	-	X	-	-
Any Hydrophobic	φ	-	-	-
Small Hydrophobic	ξ	-	-	-

(*taken from The association of Biomolecular Resource Centers:

www.abrf.org)

Publications

1. Chen, V.C.; Chou, C.-C. Hsieh, H.-Y., Perreault, H.; Khoo, K.-H. (2007) Targeted identification of phosphorylated peptides by Off-line HPLC-MALDI-MS/MS using inclusion lists based on *m/z* and retention time. *In preparation*.
2. Lattova, E.; Chen, V.C.; Varma, S.; Bezaheh, T.; Perreault, H. (2007) Matrix-assisted laser desorption/ionization on-target method for the investigation of oligosaccharides and glycosylation sites in glycopeptides and glycoproteins. *Rapid Communication Mass Spectrometry*, 21, 1644-1650.
3. Chen, Vincent C. and Perreault, H el ene (2007) Title: "Method and apparatus for depositing samples on a target surface" United States & Canada, Patent.
4. Chen, V.C.; Li, X.; Perreault, H.; Nagy, J.I. (2006) Interaction of zonula occludens-1 (ZO-1) with α -actinin-4: application of functional proteomics for identification of PDZ domain-associated proteins. *Journal of Proteome Research*, 9, 2123-2134.
5. Snovida, S.I.; Chen, V.C.; Perreault, H. (2006) Use of a novel DHB/aniline MALDI matrix for improved detection and on-target derivatization of glycans: A preliminary report. *Analytical Chemistry*, 78, 8561-8568.
6. Snovida, S.I.; Chen, V.C.; Krokhin, O.; Perreault, H. (2006) Isolation and identification of Sialylated glycopeptides from bovine α 1-acid glycoprotein by off-line capillary electrophoresis MALDI-TOF Mass Spectrometry. *Analytical Chemistry*, 78, 6556-65563.
7. Chen, V.C.; Cheng, K.; Ens, W.; Nagy, J.I.; and Perreault, H. (2004) Device for the reverse phase separation and on-target deposition of peptides incorporating a novel hydrophobic sample barrier for MALDI mass spectrometry. *Analytical Chemistry*, 76, 1189-1196.
8. Pham, V.; Zhang, W.; Chen, V.C.; Whitney, T.; Yao, J.; Froese, D.; Friesen, A.D.; Diakur, J.M.; Haque, W. (2004) Design and synthesis of novel pyridoxine 5'-phosphonates as potential anti-ischemic agents. *Journal of Medicinal Chemistry*, 46, 3680-3687.

Part 1: General Introduction

1.0 General Introduction

The ultimate aim of the biological sciences is to describe the principles that govern life. In 2001, a significant step towards reaching this goal was realized when the final sequences of the human genome were completed (Lander et al., 2001; Venter et al., 2001). With the availability of large-scale genomic data, the focus within the area of cell biology shifted from the acquisition of DNA sequences to the extraction of functional information (Bateman et al., 2002; Guigo et al., 2006, Pennisi, 2007). The cell is a complex system containing a myriad of mechanisms needed to control responses to internal and external pressures. Signals from growth factors, ligands, extracellular matrix, DNA damage, electrical excitation, shear stress, cell-cycle checkpoints, oxygen and nutrient status are continually monitored and addressed by the regulation of gene-products. The central dogma of molecular biology (Figure 1.1) outlines the flow of information from DNA, to mRNA (also known as transcripts), to polyamino acid sequences (proteins). At the protein level, introduction of post-translational modifications (PTMs) serve to further fine-tune a number of processes including protein localization, turnover and the establishment of protein-protein interactions. While knowledge of a genome can serve as a 'rough guide', its static nature does not adequately reflect the dynamic nature of any given biological system.

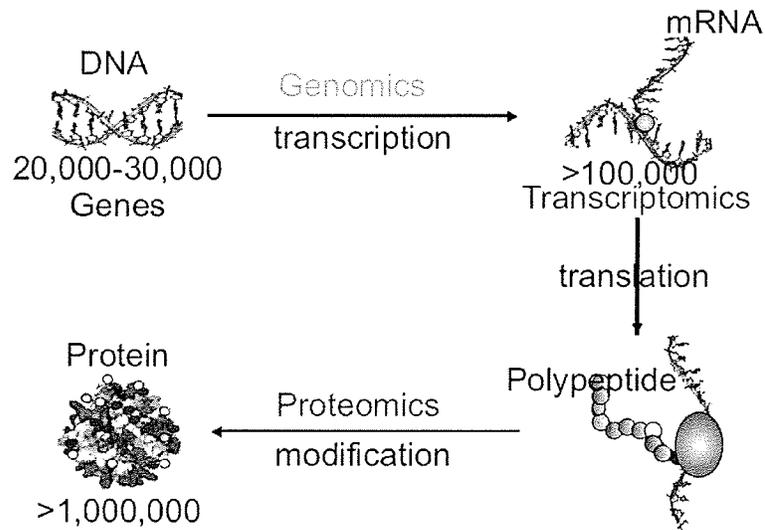


Figure 1.1: Central Dogma of Molecular Biology. Genomic and transcriptomic methods focus on determining sequences and abundances of DNA and RNA respectively while proteomic approaches are directed at determining protein abundances, modifications, cellular expression and protein-protein interactions. Analysis of the human genome has suggested anywhere from 20,000-30,000 genes are present (Pennisi, 2007). Higher levels of complexity are introduced as RNA splicing and editing can lead to more than 100,000 mRNA transcripts capable of producing protein sequences. Further, the introduction of post-translational modifications, such as phosphorylation and glycosylation, are capable of producing more than 1,000,000 distinct protein species (Stults, 2005).

Fortunately, significant advances in the experimental evaluation of gene-products and capabilities of high-throughput technologies have accelerated a number of post-genomic disciplines better suited to fill this role. Most notably, transcriptomics (analysis of RNA), and the topic of this thesis, proteomics (analysis of proteins), have led to a major paradigm shift in the way biological research is viewed and conducted. Rather than addressing a narrowly defined hypothesis, both fields attempt to decipher biology by querying large amounts of experimentally-derived data.

1.1 Transcriptomics

The total complement of mRNA in a cell or tissue at any given time constitutes a transcriptome. In eukaryotic cells, mRNA sequences bound to multiple ribosomes called polysomes undergo active translation resulting in protein synthesis, while translationally inactive mRNAs are associated with single ribosomes (monosomes). Constant flux between these two states plays a key role in controlling expression and quantities of protein synthesized. Typically achieved by cDNA microarrays, transcriptomic profiling provides insights into the expression properties of a cell or tissue (Ross, 1995; Guhaniyogi et al., 2001) and allows for the simultaneous measurement of mRNA levels for thousands of genes (Gerhold et al., 1999). High-throughput, high sensitivity and the ability to amplify low abundant transcripts (by polymerase chain reaction, PCR) have made cDNA microarrays extremely powerful tools. For example, the differential analysis of normal vs. diseased

states using cDNA microarrays has proven to be valuable for the identification of genes linked to a specific physiological condition (Cekan, 2004).

Despite the obvious attraction of profiling mRNAs, transcript-level analyses have a number of practical limitations. While transcriptomics offers the advantage of high-throughput, sensitivity and low cost, it unfortunately does not offer the capability to detect important changes in protein PTMs, protein-protein interactions and splice variants that affect the majority of proteins (Landers et al., 2001). Additionally, the use of cDNA microarray technologies is predicated on the hypothesis that protein levels within the cell are directly correlated to transcript levels. While this relationship has been demonstrated (Celis et al., 2000), a number of studies have suggested that measurement of cellular mRNAs is a poor predictor of protein levels *in vivo* (Gygi et al., 1999; Anderson et al., 1997). While mRNAs do form a template for protein synthesis, poor correlations are observed because proteins are actively regulated by a number of processes, including organelle accumulation, translocation, proteolytic processing, PTM and degradation. Furthermore, transcript profiling is further hindered by the limited types of experiments that can be performed and are often plagued by high background/noise (Hegde et al., 2003). Although more laborious than transcript-level analyses, proteomics is attractive as the functional units of the cell (i.e. proteins) are the direct subject of study. Further, as any 'front end' sampling technique resulting in a protein sample can be coupled to

proteomics, the scientific discipline has proven to be an extraordinary investigative tool capable of creatively addressing a broad range of biological questions.

1.2 Proteomics

As the primary functional units of the cell, proteins are required for an extensive range of cellular processes (Pandey et al., 2000). First coined by Marc Wilkins to mean the ‘**protein complement encoded by a genome**’ (Wilkins et al., 1996), proteomics has been recently been re-defined as:

“The study of all proteins expressed at a given point in time under given circumstances by a specific cell or tissue type [or]...the measurement of one or more protein populations or subsets found within an organism, tissue, cell, or sub-cellular compartment. A population may be defined as a set of proteins 1) within a protein-protein interaction complex; 2) linked to a specific biological process such as a signal transduction event or; 3) having a specific set of post-translational modifications such as phosphorylation” (Stults et al., 2005).

Although the theoretical premise of proteomics is fairly straightforward, in practice, the cell-wide characterization of proteins is fraught with a number of technical challenges not faced by other “-omic” disciplines. In terms of sensitivity, proteomics has largely lagged the high-throughput capabilities of microarray technologies due to the absence of comparable genetic-level amplification and sequencing methods. Furthermore, in contrast to nucleic acids, proteins are composed of a larger subset of building blocks (20 amino acids) with a much wider set of physiochemical properties. While an assorted

set of 'starting materials' are needed to produce proteins of varying function, protein diversity prevents the development of universally applicable sampling techniques (e.g. large integral membrane proteins differ significantly from small cytoplasmic molecules). To complicate the matter further, the introduction of post-translational modifications (PTM), such as protein phosphorylation, can generate a complex variety of proteins each having distinct solubilities, molecular weights, conformations and isoelectric points. In addition, actual amounts of proteins within a cell can range from a few molecules to several million copies which continuously test analytical limits and instrumental dynamic range. Despite their restrictive nature, these problems are rapidly diminishing as new proteomic technologies are continuously being developed and established ones improved. While the scope of proteomics is broad, a typical experiment can be divided into one of four categories: (1) Protein identification, (2) Protein quantification or differential analysis, (3) Protein-protein interactions, and (4) Post-translational modifications (Stults et al., 2005).

1.2.1 Protein identification.

Identification of proteins by mass spectrometry (MS) is the category that is most closely identified with proteomics. The availability of complete genome sequences has made protein identifications using MS data a common enterprise. A match between peptides generated by proteolytic cleavage (most commonly performed by trypsin) with those predicted from

theoretical digestion of database proteins (i.e. peptide-mass fingerprinting, PMF) is considered a workhorse for many proteomic endeavors. When PMF fails to provide identity, tandem mass spectrometry (MS/MS) can be used for database searches. Under some circumstances MS/MS of a single peptide can be used to identify a protein provided sufficient information such as amino acid length, redundancy, species of origin and MW are known. These elements of protein identification are fundamental to most proteomic endeavors and are discussed in greater detail in following sections.

1.2.2 Protein Quantification

Relative protein quantification is used to compare the amounts between two or more samples that may be used to determine biological differences that correlate with developmental state, drug treatment, disease or cell cycle. The ability to quantify amounts of proteins by mass spectrometry is attractive as it offers the potential to identify disease-related markers originating from any biological fluid, tissue or cell grown in culture. In its traditional application, differential analysis was performed by 2D gel electrophoresis with protein stains used to determine relative abundance. Recently, the introduction of metabolic labeling techniques such as Stable Isotopic Labeling with Amino Acids in Cell Culture (SILAC), made popular by the Mattias Mann Laboratory (Ong et al., 2002), has attracted widespread interest (Figure 1.2) (Ong et al., 2003^a; Ong et al., 2003^b; Ong et al., 2005). The principle of SILAC is fairly straightforward involving the growth of two cell

populations: one in a medium containing a 'heavy' form of an amino acid (e.g. arginine comprised of ^2H instead of ^1H , ^{13}C instead of ^{12}C , or ^{15}N instead of ^{14}N), and the second 'light' population grown in a medium of natural abundance atoms. The cells from both populations are then lysed, digested and mixed together into a single sample prior to LC-MS analysis. The metabolic incorporation of the isotopically labeled amino acid(s) results in an expected mass shift for each peptide and quantities of proteins within each population are determined by comparing their relative ion abundances. In other words, a peptide metabolically labeled with a single $^{13}\text{C}_6\text{-Arg}$ will lead to a mass increase of 6Da (experimental) compared to the peptide containing the Arg of natural abundance (control). While an intensity ratio measurement of 1:1 would indicate the expression levels are identical, a ratio of 1:3 (or more) would suggest the protein was differentially regulated due to the introduced experimental condition (e.g. disease state, drug treatment, etc.). The SILAC protocol has been applied to a number of studies including the differential analysis of prostate cancer (Everley et al., 2004; Everley et al., 2006) and the characterization of breast cancer expression profiles (Gehrmann et al., 2004).

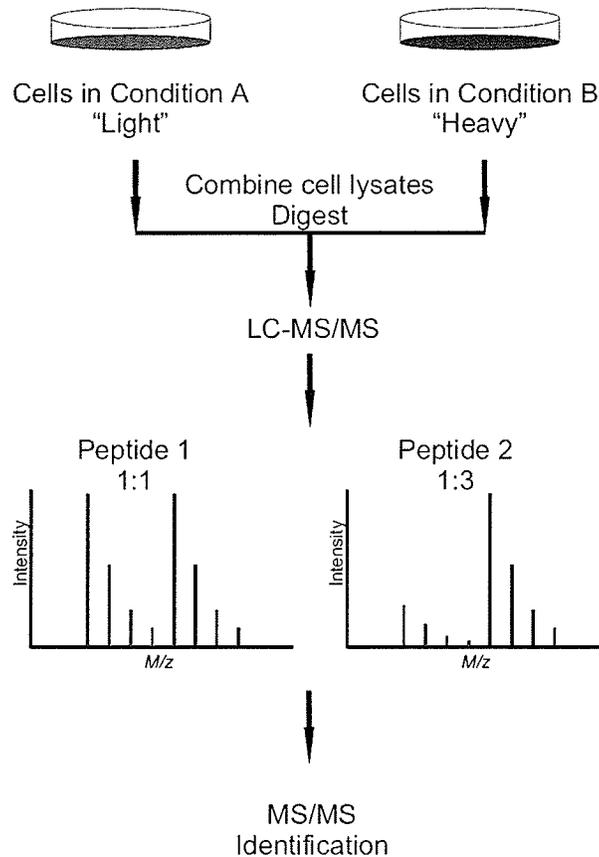


Figure 1.2: Technique for the relative quantification of protein populations by SILAC. Stably labeled amino acid in cell-culture medium is incorporated into the proteome of one cell population (Condition B, blue). Relative quantification experiments can be carried out using cells that were grown in normal medium as the control state (Condition A, red). Ratios provide relative quantities of proteins between the two protein populations. Peptide 1 (protein 1) demonstrates no change, while Peptide 2 (protein 2) becomes the subject for further investigation based on differential analysis of isotopic abundances.

1.2.3 Protein-protein interactions and functional proteomics

With the aberrant protein-protein interactions being implicated in several human diseases (Hershberger et al., 2007), it has become apparent that proteins require collaborating molecules to properly function (Pandey et al., 2000^a). The systematic characterization of protein-protein interactions is the area of proteomics which holds the greatest potential to the field of biology (Futschik et al., 2007). Unlike 'classical' applications of proteomics, this area of research, termed functional proteomics, utilizes the combined tools of molecular biology and mass spectrometry to uncover the roles of proteins by placing them with their biological partners. While this combination does not attempt to fully characterize the entire contents of the proteome, it does offer a powerful approach to probe gene-product function in both hypothesis-driven and hypothesis-generating research environments. In contrast to alternatives such as yeast-two-hybrid assays which provide the capacity to identify binary interactions, proteomic-based approaches also offer the capacity to identify interactions dependent on PTM, complexes of two or more partners and interactions of membrane proteins. Even if the limitations of yeast-hybrid systems could be tolerated, the analyses of known protein-protein interactions by yeast-two-hybrid experiments are known to generate uncomfortably high false negative rates (von Mering et al., 2002).

A recent theme in cell signal transduction is the formation of protein complexes that are regulated by the activity of modular protein-binding

elements (Skeet et al., 2006). The term 'signaling adaptor' or 'scaffolding protein', is commonly used to describe proteins entirely composed of such protein-interaction domains. The human proteome contains a number of interaction domains that collectively recognize modified and unmodified sequences. Specific examples of well known protein-protein interaction modules include Src-homology 3 (SH3) domains, SH2 and phosphotyrosine binding (PTB) domains and PDZs which bind absolute C-terminal ends (discussed further). Rather than allowing the free diffusion of signaling elements, scaffolding proteins increase the efficiency of signaling proteins by concentrating them at precise cellular locations (Zhang et al., 2003; Gimona 2006). For example, CRK is a key regulatory/scaffolding protein found at focal adhesions having a number of protein binding partners including Cas, several tyrosine kinases and small G proteins. More specifically, the SH2 domain of CRK has been shown to bind one of fifteen Cas YXXP motifs after the tyrosines undergo phosphorylation by closely associated kinases (Feller et al., 1994; Rosen et al., 1995; Donaldson et al., 2002; Feller, 2001, Takino et al., 2003; Defilippi et al., 2006). Interestingly, the mechanical deformation of cells has been shown to increase Cas tyrosine phosphorylation. Although activation of kinases in response to deformation is plausible, experimental evidence has demonstrated that the YXXP motifs are blocked by the protein's conformation in the resting cell (Sawada et al. 2006). However, once cell deformation occurs, YXXP sites of Cas become accessible, leading to their phosphorylation and the recruitment of CRK and CRK-associated proteins. In

this manner, the modular recruitment domains, binding motifs, and specific features within the CRK and Cas proteins act together to coordinate a mechanical force sensor for cell signal transduction (Pawson, 2007).

1.2.3.1 PDZ-domains

Once termed disc-large homology regions (DHRs), or GLGFs (after an originally identified glycine-leucine-glycine-phenylalanine consensus sequence), these interaction domains (Figure 1.3A) are now known by the acronym PDZ (after the first three proteins the domains were initially identified in postsynaptic protein PSD-95 (Kim et al., 1995), *Drosophila* Discs-large (Cho et al., 1992) and tight junction protein ZO-1 (Woods et al., 1993)). While they do not have membrane spanning regions, PDZs are known to localize with components of the plasma membrane. Interestingly, of the proteins that do contain PDZs, a significant proportion contain multiple copies of the domain, along with other (e.g. SH2, SH3) interaction modules. For instance, ZO-1 (Figure 1.3B) contains 3 PDZs, a SH3 (Src homology 3), a Guk (guanlyate kinase), and a 'C-terminal' ABD (actin binding domain). Although the binding partners at each site of ZO-1 have not been fully deduced, it is widely believed these domains of ZO-1 participate in the organization of macromolecular signaling complexes at tight junctions and gap junctions (Giepmans et al, 1998; Kausala et al, 2001; Itoh et al., 1993; Van Zeijl et al., 2007; Mitic et al., 1999; Fanning et al, 2002; Li et al., 2004; Balda et al., 2000).

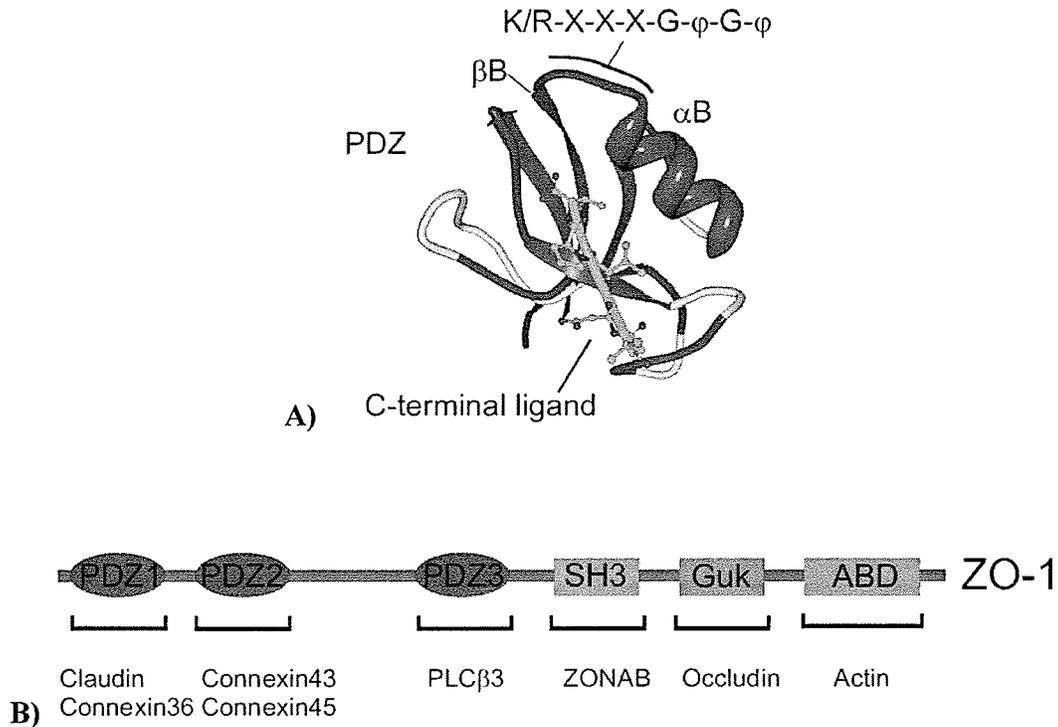


Figure 1.3: PDZs and PDZ-domain containing protein ZO-1. A) Ribbon diagram of the Erbin PDZ domain bound to the carboxyl-terminal tail of the ErbB2 receptor. The peptide ligand, shown here in green, lies in a groove created by the β B strand and α B helix. The “K/R-X-X-X-G- ϕ -G- ϕ ” carboxy-binding loop is highly conserved in PDZ modules and is responsible for recognizing and binding C-termini of target proteins. Once bound, the negatively charged carboxylic end of a targeted protein is stabilized by positively charged K/R (Lys/Arg) side chains within the carboxylate-binding loop. The hydrophobic residues (ϕ) within the loop preferentially bind ligands having hydrophobic residues located at the absolute C-terminal position (P_0) (Tochio et al., 2000; Zhang et al., 2003). This PDZ image was created in PDB Workshop 1.5 using Protein Data Bank coordinates: 1MFL for Erbin complexed with ErbB2 C-terminal peptide (Birrane et al., 2003). X represents any amino acid. B) Scaffolding protein ZO-1 showing the multiple protein-protein interaction domains and binding location of known interactors.

The primary function of the PDZ domain was originally established when the PDZ of PSD-95 was shown to bind the very C-terminal tail of Shaker-type K⁺ channels (Kim et al., 1995). Since then, a number of PDZ domains structures alone and with their respective ligands (Table 1.1) have been deduced which has furthered our understanding of the domain's recognition and molecular selectivity (Birrane et al., 2003, Songyang et al., 1999). Even though subtle variations exist, all PDZ domains have a signature K/R-X-X-X-G-φ-G/ξ-φ (where X is any amino acid, φ is a hydrophobic residue, and ξ is a small hydrophobic residue - e.g. Ala) "carboxylate binding loop" which is responsible for recognition and ligand binding (Doyle et al., 1996; Morais et al., 1996; Schultz et al., 2000, Bezprozvanny et al., 2001; Bhattacharyya et al., 2006; Ponting, 1997). The nomenclature for PDZ ligands designates the C-terminal amino acid as position 0 (P₀), with adjacent residues assigned as P₋₁, P₋₂, P₋₃ etc (from C-terminal end to N-terminus). Structures of PDZ domains with their ligands have shown that P₀ and P₋₂ positions have the most contact with the PDZ binding pocket causing sequence specific-recognition to occur through these sites. More specifically, it has been learned that the hydrophobic residues within the carboxy-binding loop are largely responsible for the preference of PDZs for hydrophobic residues at P₀, while selectivity at P₋₂ is dictated by the first residue of the PDZ's αB helix (αB1). For example, in class I PDZ domains the αB1 residue is a histidine, which preferentially recognizes serine/threonine residues in the P₋₂ position (via N-3 nitrogen of His and the hydroxyl group of Ser/Thr). In

Table 1.1: Examples of PDZ domains classified by C-terminal ligands.

PDZ Class	Target C-terminal Sequence	Protein Ligand	PDZ Containing Protein
Class I -X-S/T-X- ϕ	-E-T-D-V	Shaker K ⁺ Channel	PSD-95 (PDZ2) Kim et al., 1995
	-E-S-D-V	NMDA receptor subunit	PSD-95 (PDZ2) Kornau et al., 1995
	-Q-S-A-V	Protein Kinase C	PICK Staudinger et al., 1997
	-N-T-Q-L	Phospholipase C β (PLC β 3)	ZO-1 (PDZ3) Van Zeijl et al., 2007
Class II -X- ϕ -X- ϕ	-D-V-P-V	ErbB2	ERBIN Birrane et al., 2003
	-D-L-E-I	Connexin43	ZO-1 (PDZ2) Gipmans et al., 1998
	-S-V-W-I	Connexin45	ZO-1 (PDZ2) Kausalya et al., 2001
	-S-A-Y-V	Connexin36	ZO-1 (PDZ1) Li et al., 2004
	-V-D-S-V	Melatonin receptor	nNOS Stricker et al., 1997
Class III -X-D/E-X- ϕ	-K-E-Y-V	Claudin	ZO-1 (PDZ1) Itoh et al., 1993

class III domains, for example the PDZ domain of nNOS contains a tyrosine at the α B1 position that prefers aspartate or glutamate residues at P₋₂ (Tochio et al., 1999; Stricker et al., 1997). While the structural features of the P₀ and P₋₂ positions have dominant roles in PDZ recognition, positions P₋₁ and P₋₃, along with upstream residues, have also been shown to fine-tune affinities (Songyang et al., 1997).

1.2.3.2 Functional proteomic approaches

The experimental determination of protein-protein interactions by proteomics is based upon the premise that native associations can be maintained in solution in a manner that preserves structure and non-covalent interactions (Downard, 2006). In many situations, a protein of interest (i.e. a bait protein) is recombinantly expressed with an affinity tag or fusion protein such as a glutathione-S-transferase tag originally described by Smith and Johnson (1988). Often referred to as "pull-down" or "fishing" experiments, the protein bait is immobilized on Sepharose (or agarose) beads and incubated with a cell lysate (or tissue homogenate) to capture physiologically relevant protein-protein interactions (Figure 1.4). Associated proteins are collected by centrifugation and subsequently washed to help remove non-specific protein-bead and non-specific protein-protein aggregates. Protein complexes can be released in a number of ways including by the addition of excess glutathione (for GST-fusion proteins), boiling in the presence of DTT and detergents (e.g. SDS-PAGE loading buffer), or selectively released with a known binding ligand which competes for binding sites within the bait protein. Once proteins are isolated, identification can be achieved by MS and/or western-blot analysis.

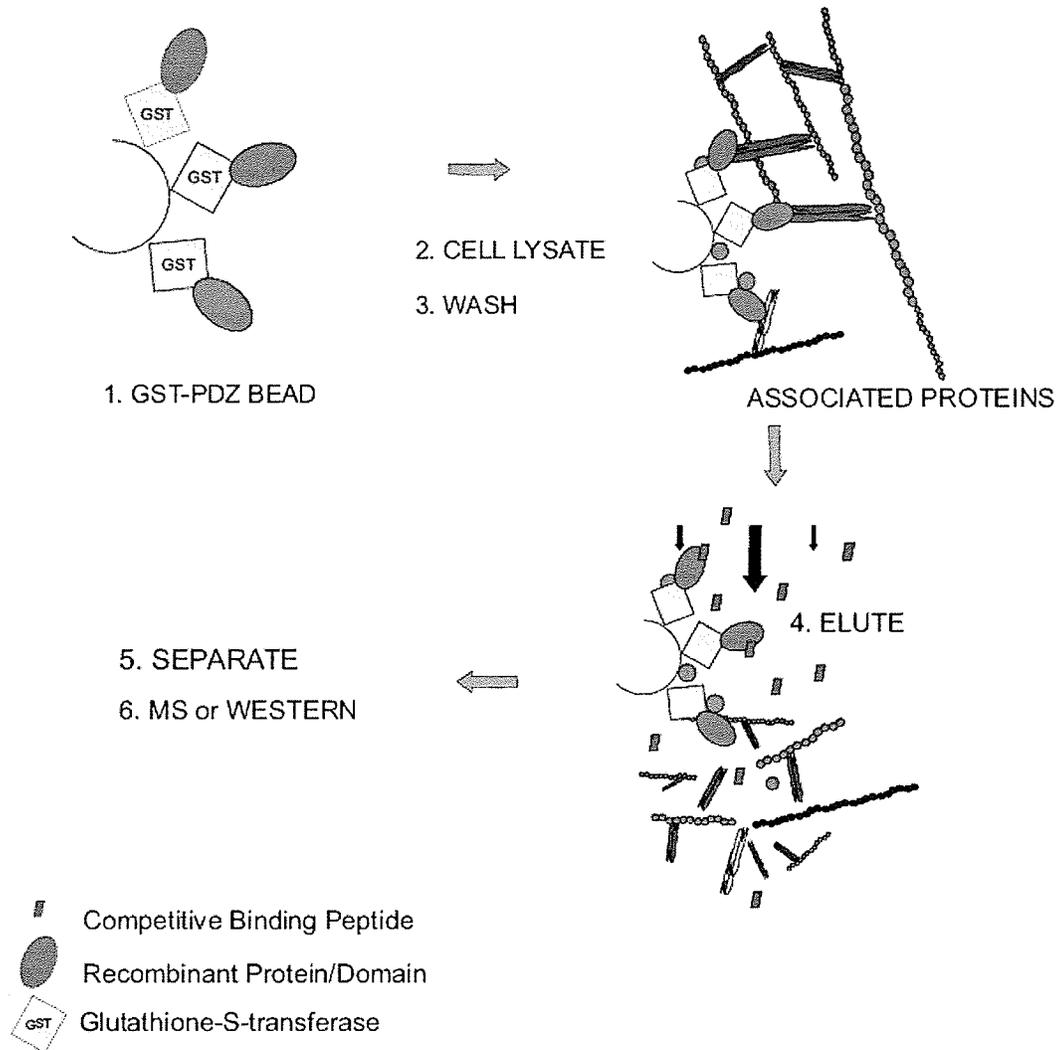


Figure 1.4: Identification of protein-protein interactions with GST-fusion proteins. (1) A recombinant GST-fusion protein containing a protein domain of interest is cross-linked to glutathione conjugated beads – GST-PDZ1 domain of ZO-1 is used here for example. (2) Cell lysates are incubated with the GST-fusion bead to establish ‘native’ associations. (3) Non-specific proteins are subsequently washed off and (4-6) bound proteins are removed by SDS-containing PAGE loading buffer or excess of a competitive (Cx36) class I PDZ C-terminal ligand as shown here.

Furthermore, as maintenance of protein structure is a prerequisite for the *in vitro* establishment of protein-protein interactions, careful selection of buffers (to closely match physiological conditions) along with steps to ensure protein bait is properly expressed. Fortunately, matrix-assisted laser desorption/ionization (MALDI) MS can be readily applied to confirm protein expression.

As an alternative to protein pull-downs, the application of immunoaffinity enrichment (immunoprecipitation, IP) is another powerful approach for the isolation of protein-protein interaction partners (Figure 1.5). The IP protocol is attractive as endogenous protein and protein complexes can be isolated and characterized (Azarkan et al., 2007). Specific examples of MS-based identifications of protein interactions include the identification of signal transduction pathways (Pandey et al., 2000^{a,b}; Jones et al., 2006; Tu et al., 2007), analysis of the interaction complex of the human spliceosome (Neubauer et al., 1998), identification of components of kinesin transport (Gindhart, 2006), and the study of the ribosomal complex and the nuclear pore complex of yeast (Link et al., 1999; Rout et al., 2000). Studies of this type have also been extended to include the analysis and identification of regulatory DNA-binding proteins (Nordhoff et al., 1999).

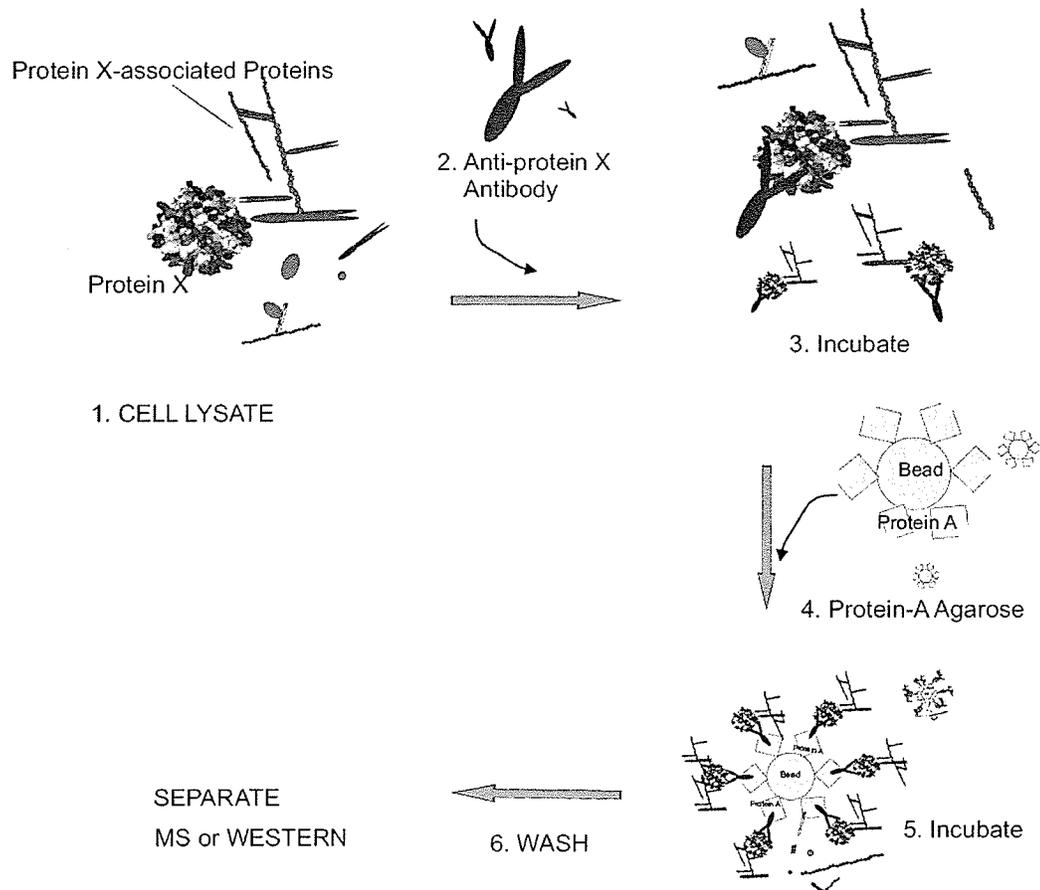


Figure 1.5: Protein-protein interaction complexes by coimmunoprecipitation. (1-3) An antibody for protein-X is incubated with cell lysates to establish the immunoprecipitation complex. (4) Associated proteins are collected using protein-A conjugated beads. (5) Protein A recognizes the Fc region of antibodies, forming a secondary non-covalent complex to pull-down protein-X and associated proteins. (6) Proteins are collected, centrifuged and washed to help remove non-specific protein aggregates for MS or Western-blot detection.

Although successful applications of functional proteomics are numerous, extensive challenges still remain. The problem with MS-based functional proteomics is that the technique is sensitive enough to detect large amounts of background protein - regardless of bait or condition used (de Hoog et al., 2004; von Mering et al., 2002). Highlighting these difficulties, a comparison of two large-scale studies despite using similar methods demonstrated a surprisingly low overlap of predicted interaction partners (Ho et al., 2002; Gavin et al., 2002). While more specific “tandem affinity” purification strategies have been developed (using probes incorporating affinity tags with enzyme specific cleavage sites), low-affinity interactions are often missed, presumably due to the increased number of sample handling steps (Rigaut et al., 1999; Azarkan et al., 2007).

1.2.4 Post-translational Modifications (PTM): Protein Phosphorylation

Protein-protein interactions are often controlled by protein post-translational modification (PTM). Currently there are more than 200 known types of protein modifications (Krishna et al, 1998). PTMs effectively expand the functional range of proteins by modifying the size, charge, structure and shape of proteins. Nevertheless, while much is known about consensus sequences, these sites of modification cannot be accurately determined from translated DNA sequences.

Protein phosphorylation is a PTM that reversibly modulates a number of important processes including inter- and intra-protein complex formation responsible for cell signal transduction (Pawson and Nash, 2000; Pawson, 2004; Skeet et al., 2006). It has been estimated that about one-third of eukaryotic proteins are phosphorylated at any given time (Zolnierowicz et al., 2000). The modification is most commonly found on hydroxyamino acids serine (S), threonine (T), and tyrosine (Y) with a population distribution of ~86-90% pS, ~10-11.8% pT and ~0.5-1.8% pY (Reinders et al., 2005; Hunter et al., 1980; Olsen et al., 2006). Phosphorylation is controlled by the dynamic activities of protein kinases and protein phosphatases (Figure 1.6) and the perturbation of these enzymes has been implicated in several forms of cancer (Robertson et al., 2000; Bonaventure et al., 2007; Alonso et al., 2004). Consequently, the experimental determination of these sites of modification is an important, yet challenging task which is not only of interest to the fields of proteomics and cell biology but also medicine.

The challenges facing researchers conducting phosphopeptide studies are considerable. Difficulties faced by phosphopeptide analyses are daunting presumably due to the low cellular concentrations of signaling molecules and the low stoichiometry of phosphopeptides relative to non-modified amino acid sequences. To complicate the matter further, a protein may be phosphorylated in a variety of ways or in a transient manner. The analysis of phosphopeptides therefore requires a minimum of two or three times more

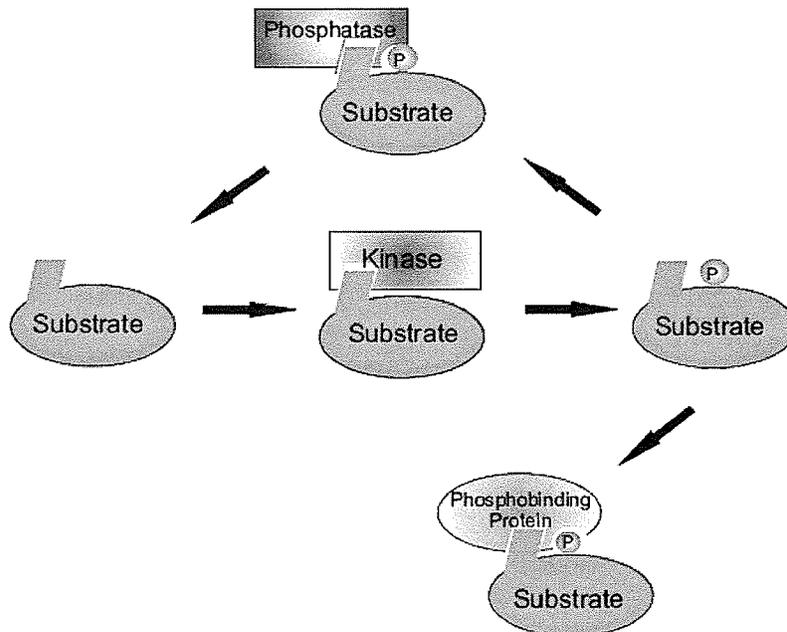


Figure 1.6: Protein Phosphorylation. A kinase recognizes and phosphorylates a protein substrate. The newly phosphorylated protein sequences can serve as a docking platform for other proteins that may be reversibly regulated by phosphatase activity.

starting material relative to what is needed to perform routine proteomic identification. However, to facilitate enrichment, a number of purification strategies have been developed taking advantage of slight variations introduced by the phosphate moiety (e.g. acidity and structure). More specifically, approaches based on affinity-enrichment and IP have proven to be the most successful (Reinders et al, 2005).

The use of phosphospecific antibodies has proven to be an elegant strategy for the enrichment phosphoproteins. Purification of phosphoproteins by IP is attractive as high quality anti-phosphotyrosine (anti-pY) antibodies are commercially available (Zheng et al., 2005; Pandey et al., 2000^c; Steen et al., 2002). Despite the successes of the anti-pY IP protocol, the applications of anti-phosphoserine (anti-pS) or anti-phosphothreonine (anti-pT) have proven somewhat more difficult - seemingly due to their lower affinities (Gronborg et al., 2002; Stannard et al., 2003).

Another recognized affinity technique for phosphoprotein/peptide enrichment is the use of immobilized affinity chromatography (IMAC). First introduced in 1961 for the purification of proteins (Hellferich, 1961), IMAC has evolved into the most widely used method for phosphopeptide purification (Hsiao et al., 2007; Nuwaysir and Stults, 1993; Posewitz and Tempst, 1999; Zhou et al. 2000; Cao and Stults, 1999). IMAC takes advantage of the natural affinity of negatively charged phosphate groups for immobilized metals Fe(III) or Ga(III). Enrichment occurs through (phosphate) non-bonding electrons coordinating to metal ions chelated to the IMAC support. The most commonly used beads for IMAC include non-porous and porous silica (El Rassi et al., 1986; McLachlin et al., 2001), agarose (Raska et al., 2002), Sepharose (Jiang et al., 1996), titanium dioxide (Larsen et al., 2005) and C18/iron nanoparticle composites (Hsiao et al., 2007). Although the technique has demonstrated utility, one of the major pitfalls of IMAC occurs from the enrichment of acidic

(Asp and Glu) peptides (Andersson et al., 1986; Mukherji et al., 2006). To overcome this problem, Ficcaro et al. (2002) introduced a method for the chemical modification (methyl esterification) of carboxylic acid side chains prior to IMAC treatment. Alternative approaches to IMAC and IP include the use of cation exchange chromatography (Beausoleil et al., 2004) and the detection of 80 Da mass shifts in spectra before and after alkaline phosphatase treatment (Larsen et al., 2001).

In terms of workflow, phosphorylation analysis by mass spectrometry is generally accomplished in a two-step approach. First, phosphoproteins of interest are proteolytically digested, ions are analyzed by MS and candidate masses are then subjected to tandem-MS (MS/MS). This may be achieved by searching for peptides corresponding to the addition of a phosphate moiety (80 Da) with MS/MS used to determine the exact position of the amino acid carrying the modification. Furthermore, phosphopeptides undergo characteristic fragmentations when subjected to MS/MS conditions, permitting them to be distinguished from their non-modified counterparts. For example, in the negative ion mode, phosphopeptides characteristically fragment to produce marker ions at m/z 79 (PO^{3-}) and 63 (PO^{2-}). MS/MS of phosphoserine- and phosphothreonine containing peptides in the positive ion mode often yield a diagnostic neutral loss of H_3PO_4 (-98 Da). These peaks ($[\text{M}+\text{nH}^+-98\text{Da}]^{n+}/n$) are often the most abundant ions in an MS/MS spectrum, but this is not always the case for phosphotyrosine residues.

Phosphotyrosines do not undergo neutral loss elimination as readily, but can produce a characteristic immonium ion at m/z 216 (Annan et al., 2001). While these methods are relatively straightforward, phosphorylated peptides typically yield lower rates of detection due to a number of factors including low sequence coverage, poor ionization, inefficient separations, and reduced recovery.

1.3 General Methodology

Proteomics aspires to deliver qualitative and quantitative descriptions of proteins within a biological system related to a given set of conditions. Although the analysis of proteins may yield important clues with respect to gene and gene-product function, they are one of the most difficult components of the cell to characterize. Even though a number of protocols exist, the majority of proteomic approaches are based on the same workflow: 1) sample preparation, 2) digestion, 3) identification followed by data interpretation. Of these steps, sample preparation is considered the most critical affecting proteomic outcome. Although problems posed are daunting, a common set of methods is available to help alleviate problems imposed by proteomic sample complexity.

Two general techniques are available for the characterization of proteomes by mass spectrometry, “bottom-up” and “top-down” (Figure 1.7). The “top-down” approach is characterized by the MS and MS/MS analysis of

whole, intact proteins while “bottom-up” approaches are characterized by digestion of sample using a proteolytic enzyme prior to mass measurement. The vast majority of mass spectrometry-based proteomics experiments, including those related to phosphoproteomics and functional proteomics, have historically taken place via the bottom-up approach. This is primarily due to the ability of mass spectrometers to provide higher quality spectral

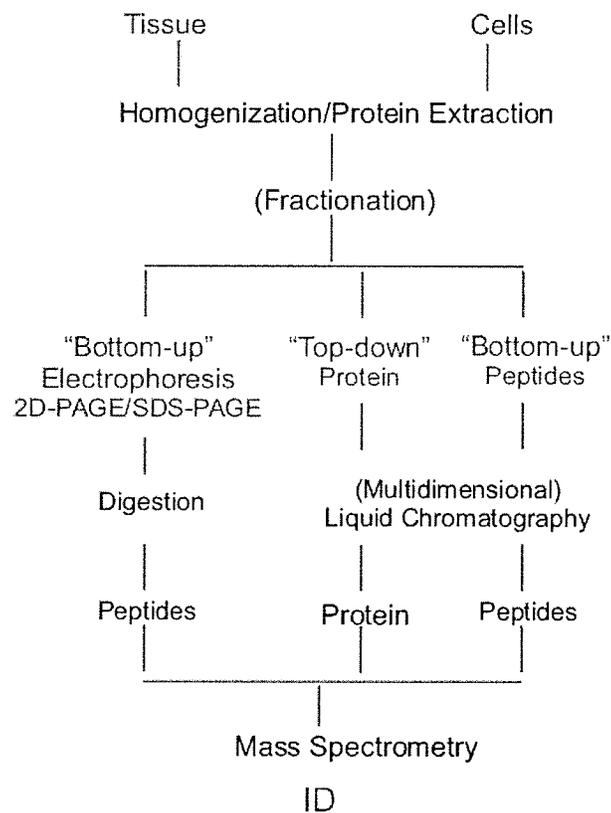


Figure 1.7: Commonly used techniques in proteomics. An overview representing some of the most frequently used separation techniques in proteomic studies, including the “bottom-up” and “top-down” approaches.

information (accuracy, precision and resolution) below m/z 4000 which is used to identify peptides (and their proteins of origin) at higher confidence levels. The bottom-up approach typically identifies a protein by a characteristic set of proteolytic peptides obtained from 1D SDS-PAGE or 2D-PAGE gels or liquid chromatography (LC). While proteomic studies have traditionally employed bottom-up workflows, recent advances in mass spectrometry have made “top-down” approaches an emerging technique, recently reviewed by Scherperel et al. (2007).

1.3.1 Gel Electrophoresis

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is one of the oldest and most highly valued multidimensional separation techniques in proteomics (O'Farrell, 1975, Wittmann-Liebold et al., 2006). The protocol employs an immobilized pH gradient (IPG) to separate proteins according to their isoelectric points (pI) in the first dimension. Proteins are then orthogonally resolved in the second dimension based on their relative electrophoretic migration within a cross-linked polyacrylamide gel. Major advantages of the 2D-PAGE protocol include high resolving power and capacity to visualize arrayed proteins after staining. Furthermore, as intensity of the stain is proportional to analyte concentration, thousands of proteins can be resolved and quantified using the 2D-PAGE format (Klose et al., 1995; Gorg et al., 2004). Because of these qualities, 2D-PAGE has been extensively used for comparative analysis of expression profiles of control vs.

experimentally perturbed proteomes. Once protein gel spots/bands are observed, they are removed and subjected to proteolytic digestion "*in gel*" (Table 8.3) to extract peptides for identification by peptide-mass fingerprinting (Patton et al., 2002).

Although offering the highest resolving power, 2D-PAGE suffers from run-to-run variations even when separating identical samples. Recently, Difference Gel Electrophoresis (DIGE) has been developed to improve comparative 2D-PAGE analyses (Tannu et al., 2006). In the DIGE method, a molecular label carrying a CyDye™ (Cy2, Cy3, or Cy5 GE Healthcare) fluorophore is used to covalently modify lysine side chain amino groups. Prior to electrophoresis, each sample and control is individually labeled with a fluorescent label having a distinct emission and excitation wavelength. Samples are then combined and run on a single 2D-PAGE gel. During the separation process, the same protein (from different samples) will migrate to the same position within the gel. Using fluorophore specific excitation/emission wavelengths, individual protein populations can be detected, allowing each protein to be individually quantified and compared within a single large format gel. This feature of the DIGE methodology greatly simplifies overlay analyses and bypasses many of the problems associated with running and comparing multiple gels.

Despite these advances, 2D-PAGE still suffers from a number of drawbacks. First, the large number of manual preparation steps and the large surface areas of exposed gel have made them susceptible to contamination. For example, adhesion of human skin (the major component of dust) is problematic as keratin will commonly interfere with analyte signals. Furthermore, because mass spectrometers are capable of detecting femtomole amounts of protein, trace amounts of dust will result in the identification of contaminants as the major sample component. The use of powder-free nitrile gloves, sleeve protectors and, if possible a HEPA filtered laminar flow hood are recommended. Even if these problems could be addressed, 2D-PAGE exhibits bias towards high abundance, highly soluble proteins and proteins having intermediate isoelectric points (pI) and molecular weights (Gygi et al., 2000; Peng et al., 2001). While proteins with low copy numbers can be resolved on larger format gels, the large quantities of starting material, along with the inability to introduce automation have posed serious challenges to proteomic studies based on the 2D-PAGE format.

1.3.2 Liquid Chromatography

A recent trend in proteomics has been towards larger-scale experiments with increasing amounts of automation. The “on-line” LC ESI-MS platform has proven to be the most popular as sample preparation, ionization and identification steps can be fully automated (Neverova et al., 2005; Issaq et al., 2005). In fact, in many laboratories, LC systems are now loaded by an

autosampler, which allows the analysis of hundreds of mixtures per day. Systems based on reverse phase (RP) LC are commonly used as high separation efficiency for peptides and proteins can be achieved and mobile phases (usually composed of acetonitrile:water) are compatible (i.e. volatile) with ESI and MALDI-type ion sources. Reverse phase LC stationary phases are generally obtained by derivatizing the surface of silica bead with alkyl chains of varying length from C4 to C18. While derivatives of C4 and C8 are generally reserved for protein separations, C18 is favored for peptide sample treatment as the smaller analytes tend to demonstrate lower hydrophobic binding capacities.

Current RP LC systems have a wide selection of chromatographic materials available to optimize peptide and protein separations as the efficiency of a LC system is largely determined by the packing material selected. Options for reverse phase packing material include particle size, pore size, surface area, stationary phase, and particle surface chemistry. The most common reverse phase column packing material is based on spherical 3 or 5 μm diameter silica particles. The separation efficiency of the packing material increases when bead diameter is reduced for the same length column. Therefore, while smaller diameter particles are highly desirable, in practice, the application of ≤ 3 μm beads tends to elevate LC back pressures beyond the capabilities of most commercial chromatographs. Additionally, increasing particle pore size from 60 to 300Å enhances substrate surface

accessibility and allows the use of higher flow rates. While increased porosity is suitable for relatively small proteins and peptides, high mass proteins with larger radii suffer reduced resolution and peak broadening due to diffusion effects (Neverova et al., 2005).

Proteome complexity often dictates that two or more purification techniques based on different mechanisms of separation are needed to adequately resolve a sample. The most recognized multidimensional approach is the 2D-PAGE platform, however, to bypass the limitations of 2D-PAGE a number of “gel-free” multidimensional peptide separation techniques have been recently developed. Although the digestion of an entire proteome greatly increases sample complexity, peptides have a narrower range of chemical properties that makes them amenable to automation using multidimensional-LC (Zhang et al., 2004). While any combination of columns with different retention properties is plausible, proteomic multidimensional LC systems are generally restricted to RP chromatography in the final dimension to maintain compatibility with MS instrumentation (Wang et al., 2003; Gilar et al., 2005).

Of the multidimensional LC techniques, the introduction of multidimensional protein identification technology (MudPit) has proven to be a powerful separation strategy (Gilar et al., 2005). Introduced by the Yates laboratory (Link et al., 1999), the MudPit protocol features a dual-layer

column composed of a strong cation exchange (SCX) resin and C18 material. To help automate 2D-LC runs, MudPit stationary phases are layered into a single column with the C18 packing located towards the tip of an ESI capillary. First dimension fractionation is performed on the SCX resin using a step gradient comprised of increasing (acetate) salts. At each SCX gradient step, a charge-based peptide fraction is released onto the C18 bed where the peptides are separated based on hydrophobicity. As ACN/H₂O gradients do not influence SCX-bound analytes, separations can be carried out analogous to traditional RP LC-MS without the use of complicated switching valves and sample loops. This cycle is repeated for each of the (SCX) gradient steps until the sample is fully depleted.

1.3.3 Principles of Mass Spectrometry and Instrumentation

The development of mass spectrometers capable of measuring biomolecules to a high degree of precision and accuracy has been credited for the rapid success of proteomics. Even though mass spectrometry was first pioneered in the beginning of the 20th century by J.J. Thomson, F.W. Aston and A.J. Dempster (Thomson, 1913; Duckworth et al., 1986), however, mass analyzers commonly associated with proteomics, such as time-of-flight, quadrupoles, ion traps and their hybrids, were developed much later during the 1950's.

Until the mid-late 1980's the fundamental problem in biological mass spectrometry was how to transfer non-volatile biological molecules into the gas phase without destroying them. Not until the introduction of soft ionization techniques such as electrospray ionization (Fenn et al., 1988) and matrix-assisted laser desorption/ionization (Karas and Hillenkamp, 1989) was this problem solved. The discovery of these ionization techniques rapidly accelerated the fields of bioanalytical mass spectrometry, proteomics and systems biology. Originally having foundations in chemistry and physics, MS provides a powerful analytical tool for scientists to interrogate biological systems. Additionally, advances in MS instrumentation, bioinformatics and the completion of large scale genome sequencing projects have led to proteomic analyses being conducted at higher levels of sensitivity and accuracy of mass detection (Ens et al., 2005; Vestal et al., 2005).

In its most fundamental application, the mass spectrometer determines the *mass-to-charge* (m/z) ratios of ions in high vacuum. In other words, mass spectrometry does not measure the mass of molecules, but instead measures their *mass-to-charge* ratios. While the MALDI process primarily produces ions with a single net charge, ESI often generates ions with multiple charge states. To determine the molecular weight of a peptide, the observed m/z value is multiplied by the charge (z) taking into account the number of charges attached to a protein (equal to z). Thus:

$$m/z = [M+nH]^{n+}/n$$

eq. 1.0

Where:

m/z = mass-to-charge ratio

M = Molecular mass of the analyte

n = number of charges

Fortunately, determining the charge state of an ion can be performed by examining the isotope peak cluster. Adjacent peaks in an isotopic cluster at 1 m/z unit difference indicate a +1 charge state, while a difference of 0.5 m/z units would indicate a charge state of +2, etc. Even though users of proteomic technologies need not perform these calculations on a regular basis, bioinformatic resources require these values for protein identification.

Mass spectrometry is a very powerful gas phase separation technique, however, it is important to understand that it is capable of analyzing only molecules carrying a net charge. While the majority of proteomic studies are carried out in positive ion mode, it should be also noted that the technique is applicable to either positive or negative polarities. In either case, for MALDI and ESI, the term ionization relates to the transfer of an analyte to the gas phase and the simultaneous acquisition of a charge from the sample environment, typically in the form of a proton (H^+) charge. Numerous types of MS analyzers are in common use in proteomic laboratories, including triple

quadrupole, ion trap, FT-MS and linear ion trap. However due to the scope of this work, only MALDI-TOF, MALDI-QqTOF and MALDI-TOF-TOF mass spectrometers and components related to their operation will be discussed as these were the principal instruments used in the present thesis work.

1.3.3.1 Matrix-assisted laser desorption/ionization (MALDI)

While at the University of Münster, Michael Karas and Franz Hillenkamp (1988) discovered that laser bombardment of crystals of a low molecular weight aromatic organic (i.e. matrix, Figure 1.8) containing a small amount of protein analyte was capable of generating intact gaseous protein ions. Termed matrix-assisted laser desorption/ionization (MALDI), a specific wavelength of light energy is absorbed by the matrix crystal and converted into heat, which causes sample vaporization (Figure 1.9). While it is not clear whether the analytes acquire their charge during the desorption process or after entering the MALDI plume, what is known, is that the sublimation of the matrix leads to the gentle analyte transfer to the gas phase.

The MALDI ion source is ideally suited for time-of-flight (TOF) mass analyzers since laser pulses serve to establish the well defined start time needed for accurate time-of-flight measurement and mass-to-charge (m/z) determination. Once formed, ions are accelerated into the TOF mass analyzer by the application of an electric potential. However as matrix sample spots need to be exposed to high vacuum (for desorption and TOF

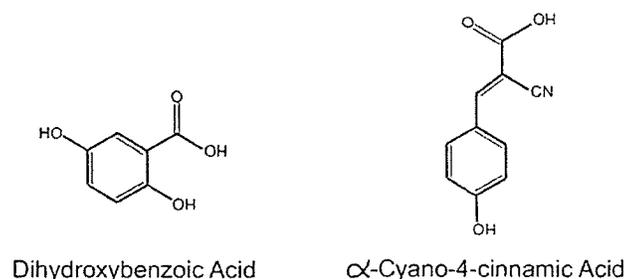


Figure 1.8: Commonly used MALDI matrices: 2,5-dihydroxybenzoic acid (DHB) and α -cyano-4 hydroxycinnamic acid (α -cyano). These are commonly used matrices used in MALDI-MS, with absorptions close to the wavelength of a N_2 laser (337 nm). MALDI matrices are usually small organic aromatic acids. A matrix solution is prepared by dissolving the organic acid in an appropriate solvent (e.g. ACN/H₂O). Peptide solutions (0.25 - 1.0 μ L) are mixed with an equal volume of matrix solution. The solution is then deposited onto a flat clean conductive MALDI surface (target), and dried to a molar ratio of 1,000 to 10,000.

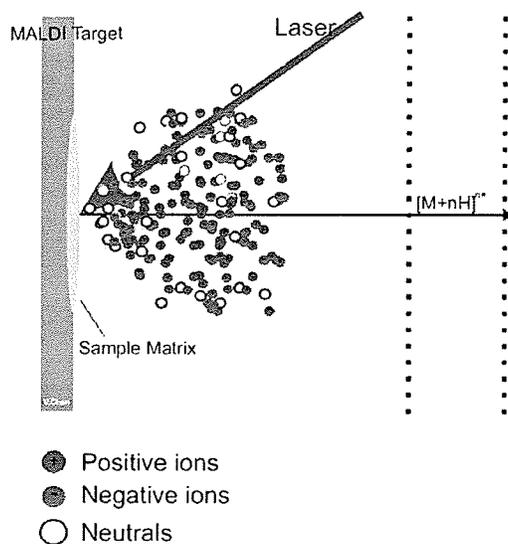


Figure 1.9: MALDI ionization process. Sample embedded in matrix is irradiated by UV laser causing localized heating and a microexplosion resulting in a plume of charged gaseous ions and neutrals.

measurement), MALDI-TOFs are not easily interfaced to on-line chromatographic separations. To bypass this problem, however, HPLC-MALDI can be performed off-line onto MALDI plates (under ambient temperatures and pressures) using a robotic deposition interface. Despite being technically demanding, the ability to store and repeatedly analyze sample spots offers distinct advantages (Krokhin et al., 2005).

1.3.3.2 Electrospray ionization (ESI)

The discovery of electrospray ionization (ESI) earned its inventor (John Fenn) a share of the Nobel Prize for Chemistry in 2002. Unlike in MALDI, the ESI process occurs under atmospheric conditions (Figure 1.10). The process is achieved by applying an electric field between the ESI capillary tip and the entrance of a mass spectrometer. The electric field induces the charged liquid at the tip of an ESI capillary to form a cone (Taylor cone). Droplets liberated from the cone undergo evaporation, shrink and explode into smaller droplets. The process continues until solvent is removed and 'bare' analyte ions remain. Moreover, while the droplets are shrinking, charge density increases, and it is therefore common to observe multiply charged analytes using the ESI source.

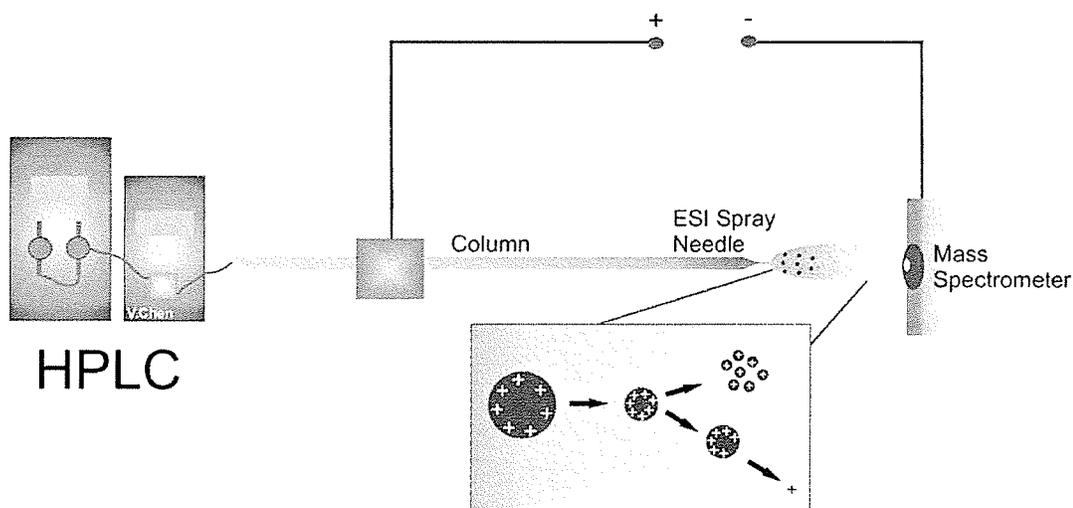


Figure 1.10: Electrospray ionization coupled to HPLC. For electrospray ionization, the tapered end of a liquid chromatography column or metal needle is held at an electrical potential of several kV with respect to the entrance (orifice) of the mass spectrometer. Once emitted, the size of charged airborne droplets continually decrease as solvent is stripped off by evaporation. Droplets undergo repetitive fission and evaporation until fully desolvated ions are produced.

1.3.3.3 Quadrupole

A quadrupole (Q) mass analyzer consists of four parallel rods. An electric field is created by applying a DC (direct current) and oscillating radio frequency (RF) voltages to opposite sets of rods, with adjacent rods having opposite polarities. As ions flow into the mass analyzer, they oscillate depending on their m/z with an amplitude dependent on the magnitude of the RF and DC voltages (Figure 1.11). The complicated motion can be modeled

mathematically using a set of equations commonly referred to as the Mathieu equations (Dawson, 1976; Dawson, 1995). For a given DC and RF potential, only ions of a given m/z will have a stable path through the mass analyzer, while others will be lost by escaping the filter or colliding with the rods. Stepping through the m/z range is achieved by applying different potentials and detecting the ions that pass through at each m/z value to generate the mass spectrum. It should be also noted that quadrupoles operating in “RF only mode” allow all ions to pass through making them ideally suited for use as ion guides (e.g. collisional cooling ion guides described later) or CID collision cells (q) for MS/MS.

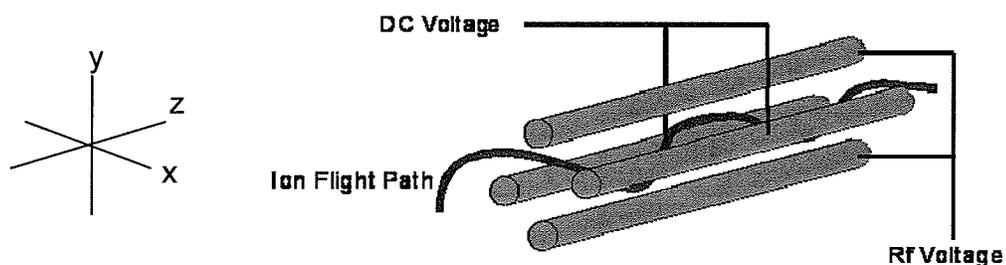


Figure 1.11: Schematic representation of a quadrupole mass filter. An ion traveling through the center of the filter (z-axis) is induced to oscillate in the xy plane. If oscillation becomes too great, ions will escape or collide with one of the rods. If however, the ion path is stable (blue), then ions are allowed to pass through the filter to the detector or subsequent sections of the mass spectrometer.

1.3.3.4 Time-of-flight

In 1955 Wiley and McLaren described the predecessor of all modern commercial time-of-flight (TOF) instruments (Wiley and McLaren, 1955). This type of mass analyzer measures the time it takes for ions to travel through a field-free flight tube. In the ion source, all the ions are given the same kinetic energy (Figure 1.12). As kinetic energy is a function of mass (equation 1.1), the lighter ions will fly faster than the heavier ones and ions will separate along the tube while moving toward the detector. Because TOF analyzers effectively measure time, instruments employing these analyzers benefit from well-defined start times afforded by a pulsed source such as a MALDI laser. Laser pulse times of 3 ns and times-of-flight of tens of microseconds are typically measured (Baldwin, 2000), making the MALDI-TOF measurements extremely fast. The high speed and low mechanical complexity of TOF instruments have made them commonly employed devices for high-throughput protein identifications.

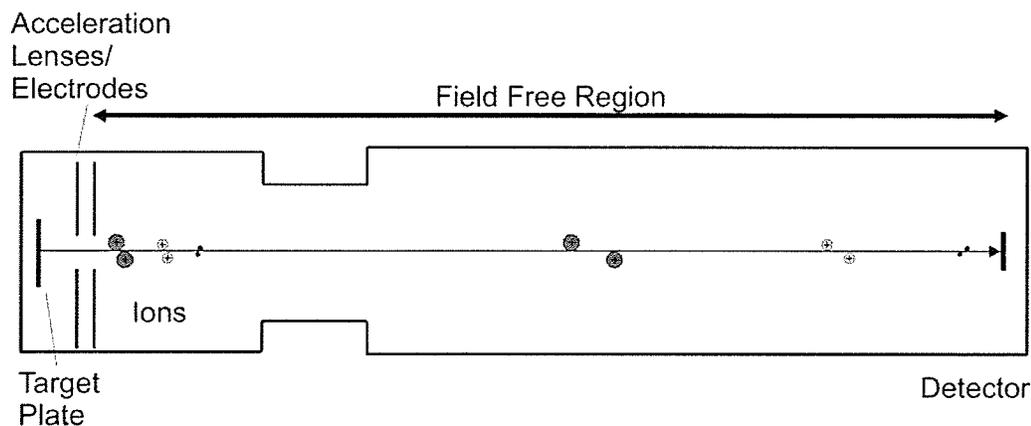


Figure 1.12: Time-of-flight mass analyzer. In its basic form, a TOF instrument includes a vacuum flight tube with an ion source at one end and a detector at the other (Wiley and McLaren, 1955). After ionization, analytes are accelerated towards a counter electrode, passing into a fixed length field-free drift tube to the detector.

The mass-to-charge (m/z) of ion can be calculated according to the following where:

KE = kinetic energy (J)

M = mass (amu)

V = velocity (m/s)

z = charge state

e = elementary unit charge ($1.6021917 \times 10^{-19}$ C)

V = applied potential (v)

For a monoenergetic packet of ions, the kinetic energy (KE) is equivalent to;

$$KE = 1/2mv^2 = zeV \quad \text{(eq. 1.1)}$$

The velocity, v , can be expressed in terms of the length of the TOF tube (L) divided by the time to traverse the tube (T_f).

$$v = L / T_f \quad \text{(eq. 1.2)}$$

Substituting 1.2 into 1.1 we obtain:

$$KE = 1/2m(L/T_f)^2 = zeV. \quad \text{(eq. 1.3)}$$

Solving for mass-to-charge (m/z) we obtain:

$$m/z = 2eV T_f^2 / L^2 \quad \text{(eq. 1.4)}$$

Within a given TOF instrument, L and e are constants, V is known, T_f is measured and thus m/z can be determined from the measured flight time.

In their basic form, TOF instruments have poor resolution due to the nature of the MALDI process - subtle variations in initial kinetic energy lead to reduced accuracy and sensitivity as they cause shifts in time-of-flight values. To compensate, all modern MALDI-TOF instruments incorporate time-lag focusing and ion mirrors to correct these variances and increase mass resolution and accuracy.

1.3.3.4.1 Time-lag focusing

In early MALDI-TOF instruments, ions were accelerated as soon as they were desorbed by the laser pulse (Figure 1.13A). This caused newly formed ions within the ion source to accelerate through a gas plume of desorbed molecules, which resulted in lower sensitivity and resolution from collisional scattering. Further losses in resolution are introduced by the explosive nature of the MALDI process, causing ions of a given mass to receive different kinetic energies, which translated into different detector impact times. To compensate, “time-lag focusing” a.k.a. “delayed extraction” (DE) was developed (Figure 1.13B) (Brown *et al*, 1995; Vestal *et al*, 1995). Instead of applying a continuous extraction field, a DE potential is applied 100-400 ns after the laser pulse is initiated. During this short delay, ions are allowed to travel distances proportional to their initial velocities. At the end of this period, the acceleration potential is applied and ions closest to the target surface (i.e. initially slower moving ions) receive a proportionally greater velocity than their (initially) faster moving counter parts. The application of the potential in this manner effectively allows all the ions of a given m/z to focus within the drift tube and impact the detector at the same time. Additionally, because ions are not rapidly pulled through the MALDI plume of neutral and oppositely charged molecules, corresponding increases in sensitivity are also observed (Vestal *et al.*, 1995).

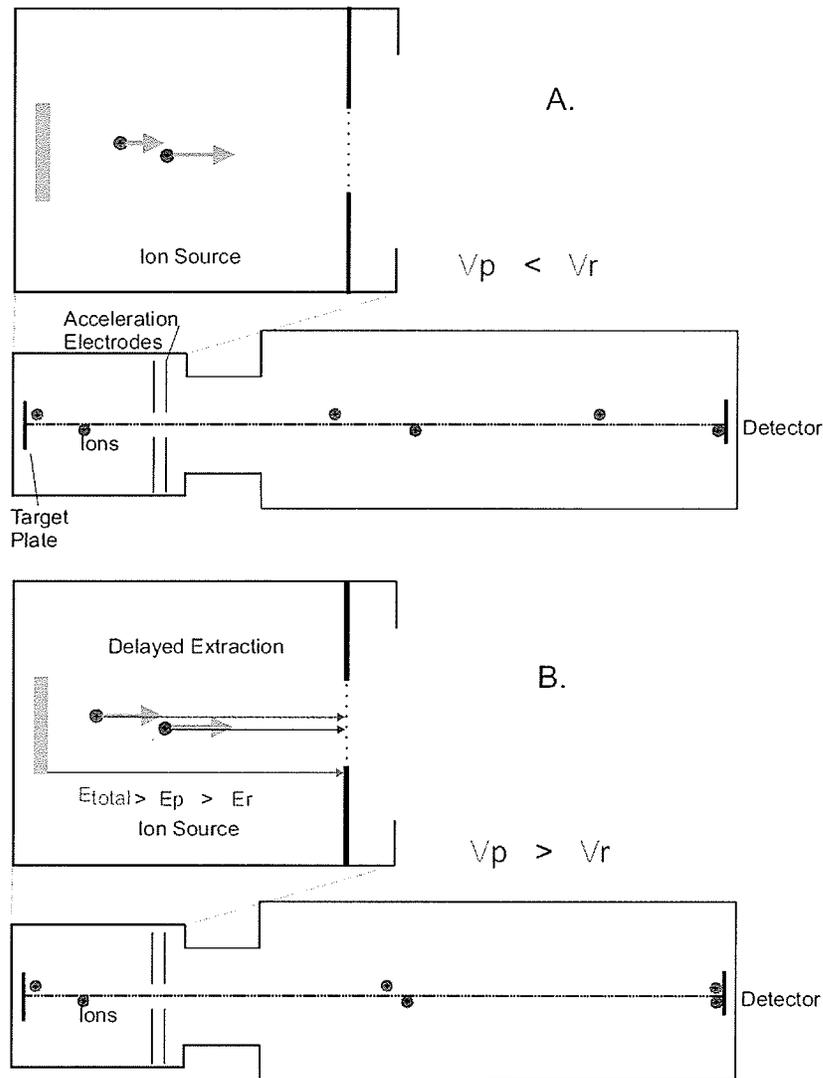


Figure 1.13: Principles of time-lag focusing. **A)** Ions R (red) and P (purple) have the same m/z but different K.E. Spatial and velocity differences imparted by the MALDI process are translated through the field free region, leading to different impact times. **B)** Principles of time lag focusing: A short time after ionization, the ion extraction field is applied. As the extraction field is delayed, slower moving ions (purple) are accelerated to a higher velocities (V_p) due to their relative proximity to the target surface (which is held at maximum acceleration potential E_{total}). This feature allows the slower moving ions to “catch-up” with their faster moving ions and focus at the detector. E, K.E. imparted to the ion; V, velocity.

1.3.3.4.2 Ion Mirror

The application of an ion mirror (Reflectron®) further enhances analyses by fine tuning small energy differences within a set of analyte ions that were not corrected by time-lag focusing (Mamyrin et al., 1979) (Figure 1.14). The Reflectron® compensates for energy variations by effectively increasing the path lengths of faster moving ions compared to slower ions. An ion mirror, fitted at one end of a linear flight tube, uses an electrostatic field to reflect ions at a small angle towards the detector. Ions carrying a higher KE will arrive earlier and will penetrate the mirror more deeply (due to increased momentum) than their more slower moving counterparts. The ion mirror improves resolving power and mass accuracy by permitting both sets of ions to exit the device at the same time.

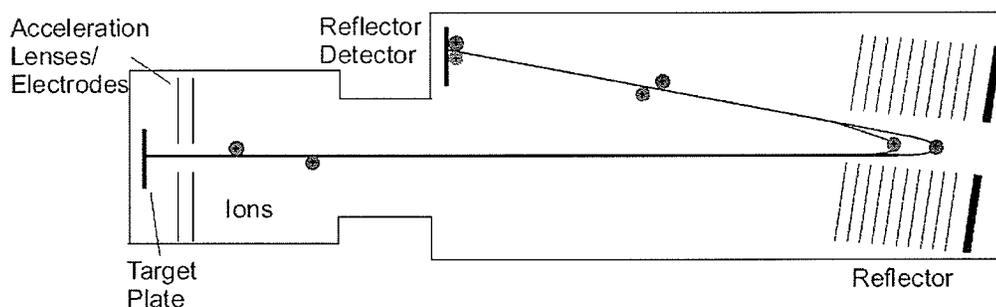


Figure 1.14: Schematic representation of a MALDI reflectron-TOF. Ions (red and purple) have the same m/z but different K.E. After ionization, analytes are accelerated towards the ion reflector where faster (higher momentum) ions (red) penetrate the mirror deeper than the slower moving ions (purple). The use of the ion mirror helps to increase resolution by effectively allowing ions to impact the detector as a well defined packet.

X/y deflectors located within the TOF are used to deflect low molecular weight ions (matrix and salts) to enhance detector sensitivity.

1.3.4.5 Matrix-assisted laser desorption/ionization quadrupole-quadrupole time-of-flight mass spectrometry (QqTOF)

In recent years MS has emerged as a major analytical tool for the identification of proteins. Although peptide-mass fingerprinting (PMF) has become the method of choice for many proteomic applications, the technique is incapable of providing the structural information that is needed to characterize peptides carrying PTMs. These limitations of single MS have made instruments capable of tandem-MS (MS/MS) especially valuable tools. In fact, an increasing number of studies have bypassed the generation of PMF data altogether, solely relying on LC-MS/MS runs to identify proteins. In a typical MS/MS experiment, a “precursor” or “parent ion” of interest is selected in the first mass spectrometer and then broken up within a “collision cell”. These impacts with the collision cell gas (N_2 , Ar) result in the production of fragment (“daughter”) ions that are measured by the final MS section for structural characterization.

The most frequently used tandem-MS instrument uses the triple quadrupole ($Q_1q_2Q_3$) configuration. In these devices, Q_1 and Q_3 serve as precursor mass selectors and for daughter ion scanning, respectively. The primary advantage of QqQ instrumentation resides in the ability to select (Q_1) and transmit ions to the (q_2) collision cell in an efficient manner. However, as

Q_3 can only permit one m/z to pass to the detector at any given moment, scanning through the mass range will result in a considerable loss of daughter ion sensitivity. In other words, although an abundant/diagnostic array of fragment ions may be produced by Q_1q_2 , the vast majority of these ions are filtered and lost during the scanning process. These limitations of triple quadrupole instruments have motivated researchers at the University of Manitoba to develop the Manitoba IV hybrid MALDI-QqTOF (Figure 1.15) (Ens et al., 2005) mass spectrometer. This MALDI instrument offers the capacity to examine the full complement of daughter ions within a single scan leading to subsequent increases in MS/MS sensitivity (Chernushevich et al., 1999).

Due to the high velocity and angular spread of MALDI ions, early attempts to measure quadrupole filtered ions by time-of-flight instruments disappointingly led to unimpressive results (Ens et al., 2005; Krutchinsky et al., 1998^a; Spengler and Cotter, 1990). A major innovation to the QqTOF configuration was realized by the addition of a collisional cooling interface. The interface is a RF quadrupole operated at a relatively high pressure (0.01-1.0 Torr) referred to as q_0 (Krutchinsky et al., 1998^a). The ion beam quality is enhanced by reducing velocity and spatial distributions by collisional dampening. The collisional cooling interface reduces velocity and spatial distributions of ions, in a manner which not only enhances the filtering

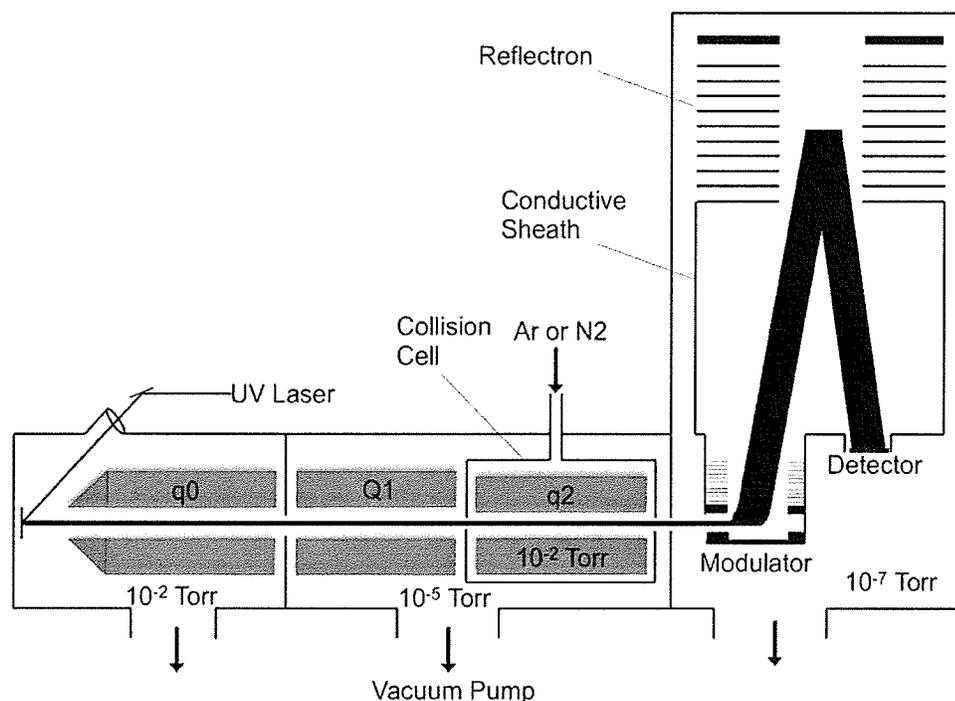


Figure 1.15: Prototype Manitoba IV/Sciex MALDI Quadrupole-quadrupole time-of-flight mass spectrometer (MALDI-QqTOF-MS). Schematic representation showing the relative positions of the collisional cooling quadrupole ion beam guide (q0), parent ion selection quadrupole (Q1), collision cell (q2) and ion mirror TOF-tube (Krutchinsky et al., 1998; Ens et al., 2005). This is the instrument used for peptide fragmentation experiments described in Parts 2 and 3.

efficiency of the (Q_1) mass selection but also provides favorable ion beam characteristics for orthogonal TOF injection (Krutchinsky et al. 1998^b).

1.3.4.6 MALDI-Time-of-flight/Time-of-flight

In 2000, a hybrid MALDI instrument (Figure 1.16), with two TOF tubes separated by a collision cell, was described (Medzihradzsky et al., 2000). The

two TOF tubes form a continuum in the MS mode as in traditional MALDI-TOF instruments. However, the main difference resides in the use of the 'timed-ion-selector' (TIS) located at the end of the first flight tube which permits for precursor ion selection in MS/MS mode (Vestal et al., 2005). The TIS ion gate is programmed to open when the lightest mass of the molecular species (monoisotope) of interest enters. The velocities of selected ions are reduced in a well defined way to impart a specific collision energy dictated by the potential difference between the source and collision cell. After fragmentation, daughter ions are analyzed in the second TOF which houses an ion mirror. Fragmentations within a TOF-TOF instrument are closely related to those occurring within an MALDI reflectron-TOF operating in post-source decay (PSD) mode (Kaufman et al., 1994) except that limitations on PSD speed, sensitivity, resolution, and mass accuracy have been overcome (Vestal et al., 2005).

The Applied Biosystems 4700 and 4800™ TOF/TOF Proteomic Analyzers are commercial versions of the instrument initially described by Medzihradszky and co-workers (2000). These instruments are equipped with a high frequency Nd:YAG laser operating at a frequency between 20 Hz – 2 kHz for rapid data generation. The 4700™ is capable of performing automated protein identifications by PMF or MS/MS fragmentation. The instrument, using the laser at 200 Hz, is capable of processing a 96 spot MALDI target in 4 minutes. Interfaced to real-time high-throughput data-base searching,

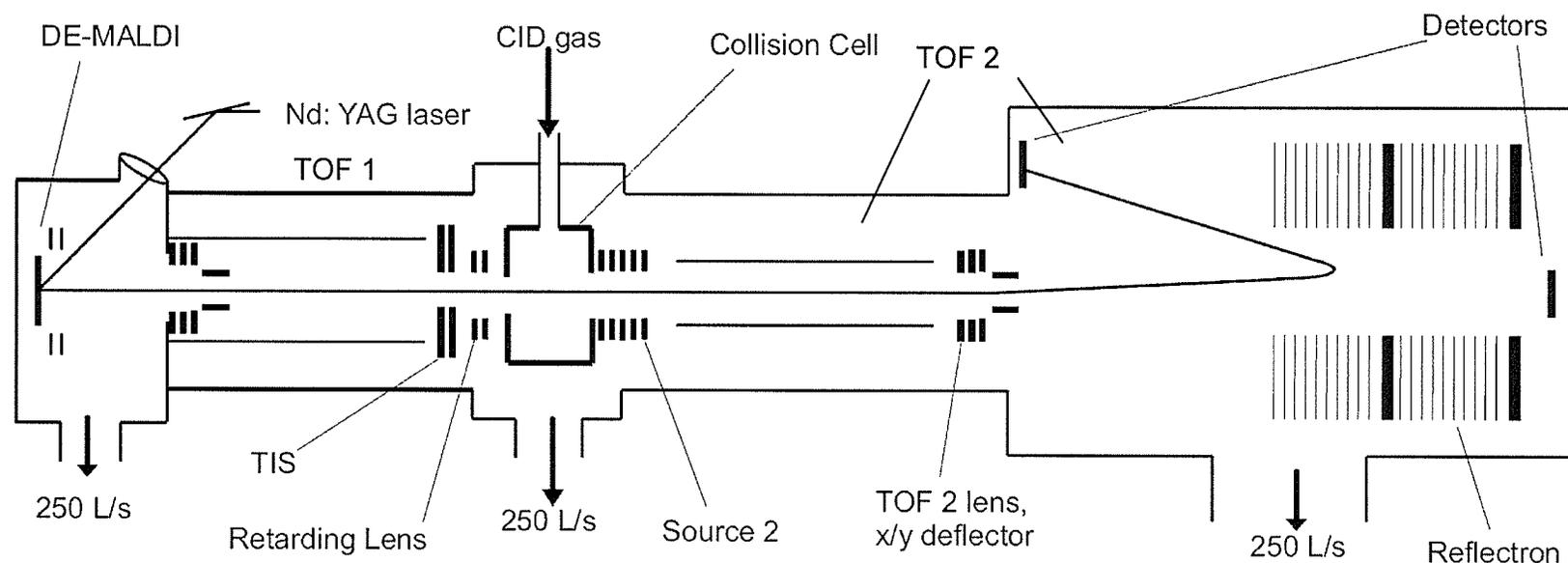


Figure 1.16: MALDI Time-of-flight/time-of-flight mass spectrometer. Schematic representation of the AB 4700™ TOF-TOF Proteomic Analyzer based on the instrumental design described by Vestal et al. (2005). The instrument is capable of MS and MS/MS modes. In MS/MS mode, the DE MALDI ion source is used to produce ions and parent ion selection is performed by the first TOF mass analyzer. Ions are focused in space and time at the timed-ion-selector (TIS) and ions outside the mass range are deflected. KE and fragmentation energy are modulated by the retarding lens and absolute potential energy difference between the source and collision cell. Parent and fragment ions proceed to the second ion source where they are reaccelerated for TOF analysis. The X/y deflector is used to reduce spectral contributions from parent ion and parent ions metastable decay (Vestal et al., 2005). This instrument was used used in Part 4.

samples can be identified as MS and MS/MS data is acquired. Furthermore, data-dependent MS/MS analysis can be conducted to determine identities of peptides in a fully automated manner.

1.3.4.7 Protein Identification by Peptide Mass-Fingerprinting (PMF)

Peptide-mass fingerprinting (PMF) is the simplest and fastest protein identification method in proteomics (Figure 1.17). PMF has become a method of choice, requiring low pmole to high femtomole amounts of protein and instrument times of only 20-30 seconds/sample (Mann et al., 1993). After measurement, monoisotopic peaks (or masses) are submitted to a database search engine which compares experimental masses to 'virtual' digests of translated genes and identities. For highly purified protein, only a small number of accurately determined masses are needed for successful identification by PMF. As a close resemblance between the experimental and the theoretical masses are used as the premise for identification, the completion of genomic sequencing projects has dramatically increased PMF success rates. However, as the PMF process attempts to match all the masses within a MS measurement, the presence of contaminant(s), artifacts from human-skin/dust (i.e. keratin) and/or peptides from two or more proteins will generally lower identification scores (Thiede et al., 2005). Likewise, the choice of search parameters will also affect identifications as the inclusion of post-translational modifications will lower the quantity of unassigned peaks, at

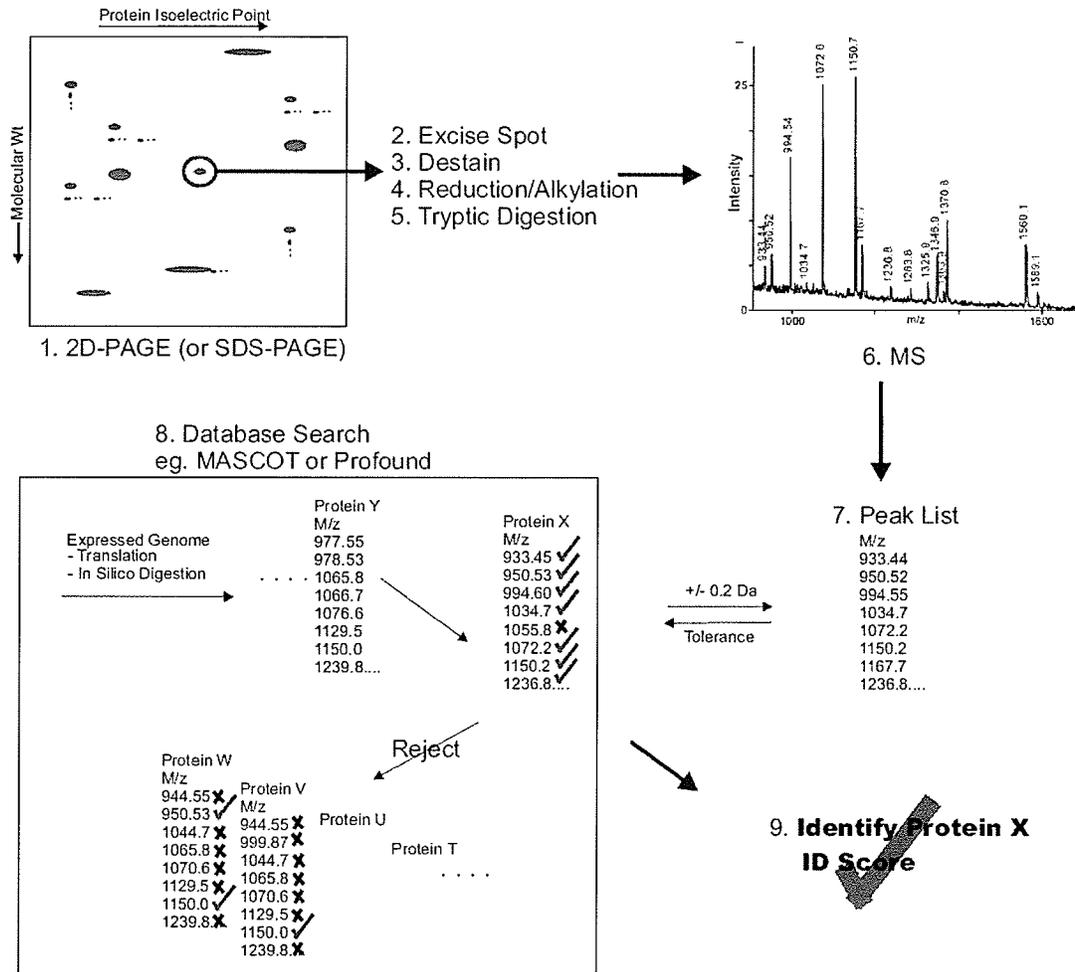


Figure 1.17: Principles of peptide-mass fingerprinting (PMF). (1) Section of the (1D or 2D) gel, corresponding to a purified protein (i.e. band/spot) of interest is removed for *in gel* digestion and MS measurement (2-5). (6,7) Experimentally derived peak lists are compared to (theoretical) *in silico* protein digests based on user-defined values including mass tolerance, proteolytic enzyme, protein Mr, chemistry (e.g. cysteine modification, chemical labeling) and PTM. Algorithms like MASCOT score identifications based on the quality and quantity of peptide matches.

the cost of generating higher rates of false positives. Furthermore, as only the mass (m/z) of an analyte is determined in single MS mode, no structural information is ultimately provided (i.e. peptide LQWSETYR will have the same mass as EIWTQYSR). This limitation effectively prohibits the proteomic identification/characterization of PTMs, isoforms and complex (e.g. shotgun) mixtures by PMF (Karty et al., 2002).

1.3.4.8 Protein Identification by Tandem-MS (MS/MS)

Although MS is typically used for the first series of measurements, situations will arise when the application of PMF does not yield a high confidence match. When this happens, fragmentation measurements by tandem-MS (MS/MS) are often required. Fragmentation spectra are most often obtained by HPLC interfaced to an ESI-based MS/MS instrument (as previously represented by Figure 1.10). Here, the products of protein digestion are separated by HPLC and as they elute they are ionized by an ESI-type (e.g. ESI, μ -spray, nanospray) source. Immediately thereafter, these components are measured in MS mode and precursor ions having signal intensity above a specified threshold are used to automatically trigger the acquisition of MS/MS spectra through a process called "data-dependent analysis". This LC-MS/MS platform has become the method of choice for many applications as sample preparation and MS introduction and protein identification are readily automated. Although suitable for many applications, the on-line LC-MS coupling does have limitations with respect to MS/MS

sensitivity. For example, under an ideal situation, every component eluting from a LC column should undergo fragmentation analysis, however in reality the limited duty cycle of mass spectrometers (the amount of time required for a MS or MS/MS scan and for switching between the two modes of operation) and the temporal restrictions introduced by the on-line LC-MS coupling, severely limit the quality of MS/MS data. These limitations of on-line HPLC-ESI are discussed in greater detail in subsequent sections.

During an MS/MS acquisition, all but the precursor ions are not allowed to enter the collision cell. Within the collision cell, ions are elevated to higher energies, which is then dissipated through a process called collision induced dissociation (CID). The analysis of MS/MS patterns over the years has led to the understanding of the mechanism of CID and a commonly used (peptide) fragment ion nomenclature (Figure 1.18) (Roepstorff et al., 1984; Johnson et al. 1987; Paizs et al., 2005). As either or both 'ends' of a fragment can carry a charge, N-terminal fragments (containing the original amino terminus) are designated a_n , b_n and c_n , while portions containing the (original) C-terminus are labeled using an x_n , y_n and z_n (where n represents the number of R-groups the fragment contains). Within a given ion series, each peak differs from its neighbor by the addition/loss of one amino acid residue. As demonstrated in Figure 1.19, the mass difference between adjacent peaks can be used to rebuild all/or a portion of the sequence and locate the position of important PTMs. Although a large number of possible fragment ions can be

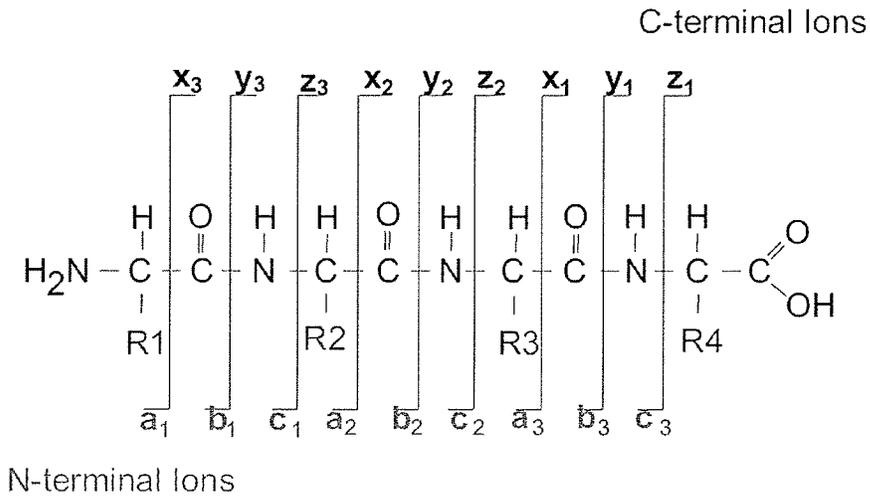


Figure 1.18: Peptide fragmentation nomenclature. To designate MS/MS peptide fragmentation, ions are consecutively labeled according to the cleavage position along the peptide backbone. For example, fragments containing the original N-terminal end within the tetrapeptide shown are consecutively designated $a_1, b_1, c_1 \dots a_3, b_3, c_3$ (indicated in green) depending on the position of the bond cleaved (as shown by long vertical lines). Alternatively, as either (or both) ends of a fragment may retain a charge, C-terminal ions can likewise be detected. Starting from the original C-terminus, fragments are consecutively labeled $x_1, y_1, z_1 \dots x_3, y_3, z_3$ (as shown in blue). Due to the preference of the fragmentation process (via the cleavage of the peptide C-N bond), B- and Y-ion series generally dominate MS/MS spectra.

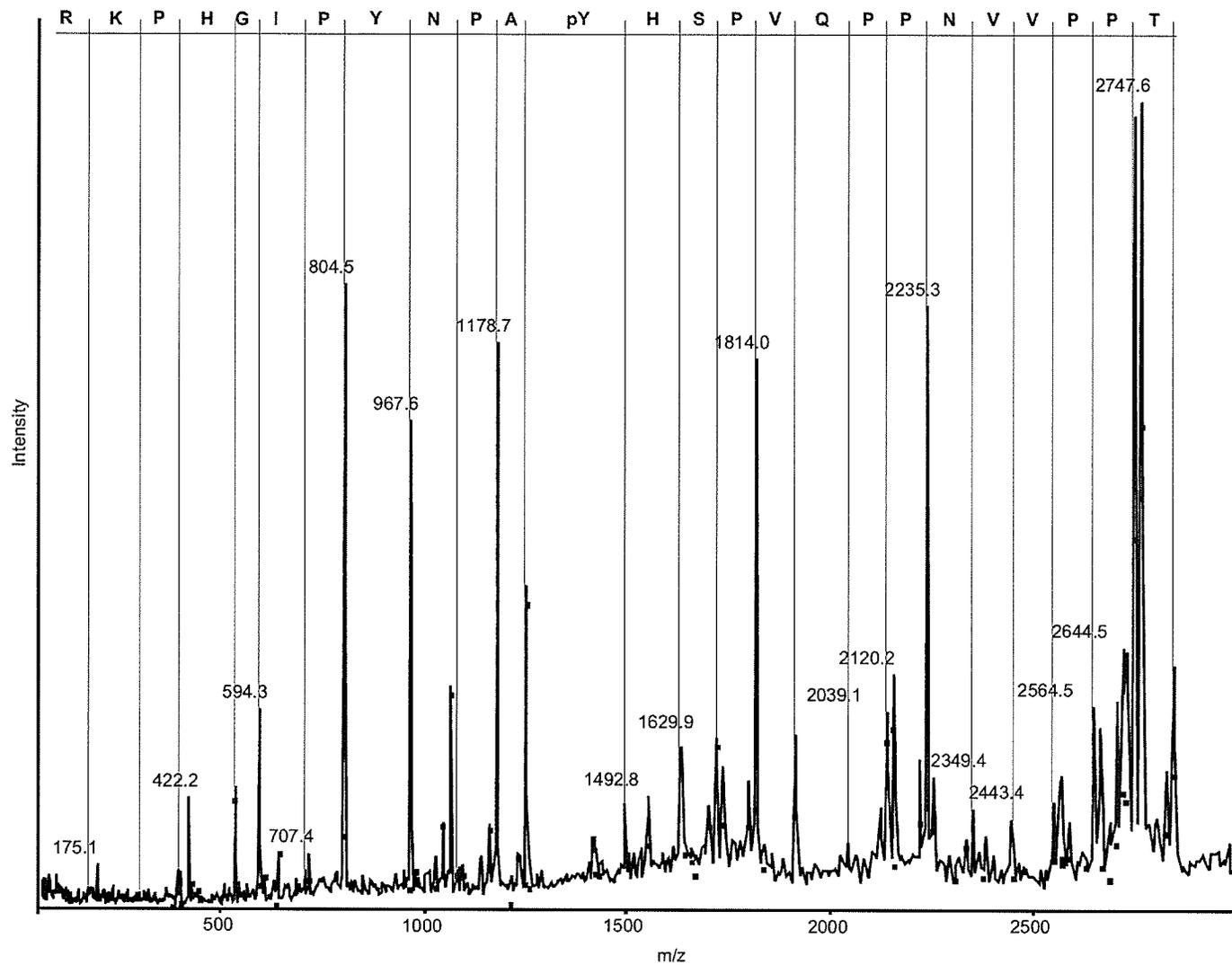


Figure 1.19: MS/MS data for the analysis of peptide sequences. Mass differences between adjacent peaks can be used to rebuild all or a portion of a peptide sequence and identify the position of a PTM (location of a phosphotyrosine (pY) shown in red). A complete set of B-ions (blue) for this peptide is observed. Peptide originated from platelet-derived growth factor receptor (PDGF-R) obtained from *Drosophila* S2 cells.

theoretically generated, the interpretation of MS/MS data is often made easier by the preference of the CID process to proceed through the lowest energy pathway, which typically corresponds to the cleavage of the peptide (C-N) bond.

In terms of protein identifications by MS/MS, like PMF, MS/MS search programs use the specificity of the protease to predict the peptide mass and then attempts to match experimentally observed ions to ones theoretically predicted within a database (Figure 1.20). Several websites offer free MS/MS data-bases including MASCOT (www.matrixscience.com), and Sonar MS/MS (www.prowl.rockefeller.edu). The approach used by MASCOT, initially described by Perkins et al. (1999), attempts to match the most intense fragment ions with the predicted B and Y-ions within a database and provides a score that is used to describe the confidence of the match. A MASCOT score between 1 and 40 indicates a moderate-to-high probability the assigned identity is random, while scores of 100 or more are considered confident (a score of ~100 corresponds to a probability of 1 in 500,000 the identification is random). Although seemingly straightforward, major complications with interpreting MS/MS spectra originate from the generation of an incomplete fragment ion series, the presence of 'undefined' peaks (which confuse automated assignments) and/or when subtle but important peaks within a spectrum are missed. For example, a MS/MS spectrum containing 2 peaks corresponding to a mass difference of 114Da could possibly correspond to

either one asparagine (114 Da) or two glycine (2 x 57 Da) residues. Although automated programs are extensively used, expert evaluations are still required at this time to validate identities having low-to-moderate scores and peptides carrying PTMs.

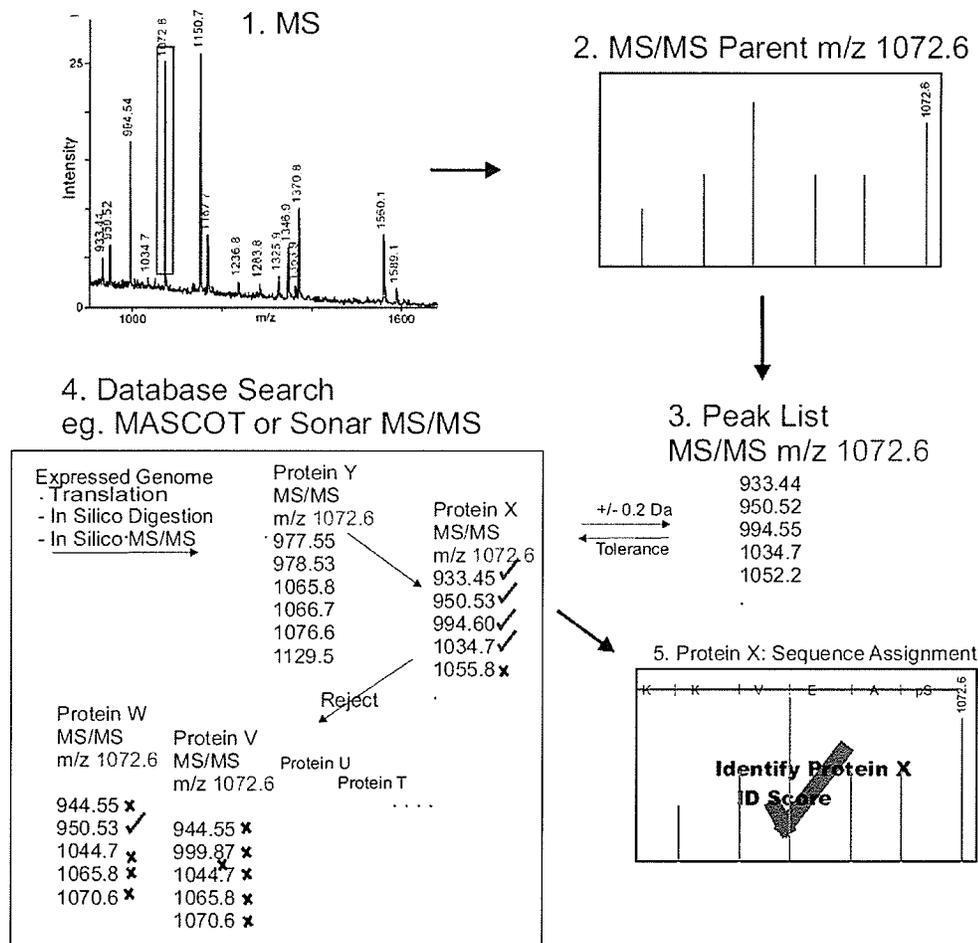


Figure 1.20: Strategy for protein identification using MS/MS data. (1) A precursor ion selected in MS mode is isolated and (2) subjected to MS/MS analysis. (3) Mass lists comprised of precursor ion mass and fragment ion masses are used as database search parameters. (4) Lists of peptide masses falling within user-defined tolerances for mass and known proteolytic enzyme and/or chemical derivatization (e.g. cysteine modification) and/or PTM are

generated. (5) Commonly employed search engines calculates expected fragment ions of each peptide within the list, while PTMs can be identified from mass shifts at specific amino acid residues.

1.4 Rationale for Studies

Cells respond to an array of stimuli by activating intracellular signaling events that involve the expression (and repression) of gene products, PTMs and protein-protein interactions. Proteomics aims to decipher these events by analyzing the total protein content within a cell or tissue-type. Despite significant advances, no proteome to date has been fully characterized. Even studies of 'simplified' proteomic subsets (e.g. organelles, interaction complexes and PTMs) have demonstrated that significant efforts are still required. Unexpected to many, the limiting factor of all proteomic studies is not related to the identification process, but rather the quality of the separation and sample handling steps taken prior to MS study (Neerova et al., 2005). In fact, due to its complexity, the proteome today cannot be resolved by a single chromatographic or electrophoretic technology. Fundamental studies in the area of the separation sciences, including those relevant to the by the field of proteomics, have continually demonstrated that the most effective strategies require the application of a number of individual separation steps such that:

"...components be subjected to two or more largely independent separation displacements. However...the separation must be structured such that whenever two components are adequately resolved in any one displacement step, they generally remain resolved throughout the process" (Gidding, 1984).

Additionally, for all the advances within the field proteomics, little ground has been gained in the development of robust microvolume approaches to one of the most frequently performed tasks within the life sciences – the purification and enrichment of peptides and proteins. Unlike DNA and RNA, there are no bioenzymatic techniques analogous to PCR that allow for the exponential amplification/purification of a protein. As a result, protein samples tend to be precious in nature and the development of any technique that provides the high performance purification of proteins at the smallest scales will be of tremendous benefit. Within this thesis are a number of strategies based on the off-line LC MALDI platform that can be readily incorporated into any multidimensional approach. Although “global-scale” proteomic analyses were not attempted, a goal of the work was to develop strategies to understand fundamental processes of the cell through the use of MALDI-mass spectrometry.

Due to the relative ease of handling smaller peptides compared to proteins, proteomics has traditionally followed bottom-up protocols. In the majority of cases, the first series of measurements are performed in MS-mode; however when this fails to yield a high confidence match, MS/MS is required. In most fragmentation experiments, peptides are separated by HPLC and infused into the emitting tip of an ESI-based ion source. While this interface does offer conveniences, the on-line approach can adversely affect peptide characterization. Even with the recent developments of

chromatographic devices with higher resolution and instrumentation with faster MS/MS acquisition rates (Wolters et al., 2001; Plumb et al., 2006; Luo et al., 2005), temporal restrictions involved with the on-line LC-MS coupling is still limited. For example, in a typical data-dependent scan, the peptides eluting from an LC column are examined in MS mode and the most intense signals are automatically selected for MS/MS (Figure 1.21). While this ion intensity-based (MS/MS) 'triggering' is suitable for the identification of higher abundance peptides, proteins of low abundance are often missed and their identities are often under represented (Picotti et al., 2007). While on-line approaches are beneficial for high-throughput analyses, the indiscriminant triggering of MS/MS spectra based on ion intensity often makes it difficult to place molecules within their functional context. While failure to distinguish a protein within a proteomic survey is possible, this failure does not indicate that the protein was not expressed, nor does it mean the protein sample was not of sufficient quality for MS/MS fragmentation. Although automated high-throughput approaches are highly valued, they generally lack the capacity to efficiently identify proteins known to be contained within a sample. While not important for many "global" proteomic applications (which aim to discover as many proteins as possible), the hypothesis-driven characterization (e.g. identification, and discovery of PTM and protein-protein interactions) of a given protein is a prerequisite to determining gene-product function.

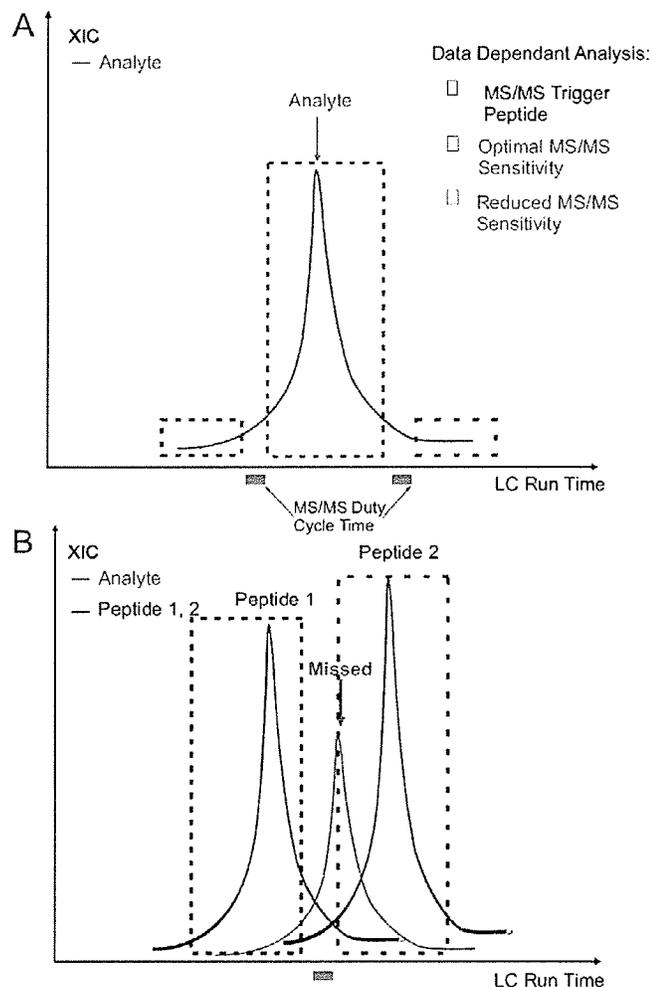


Figure 1.21: Data-dependent analysis by on-line LC-MS/MS. (A) Extracted ion chromatogram (XIC) showing the highest intensity ion (eluting from an LC run). This ion can be subjected to MS/MS analysis at maximum peak height (red rectangle) or at lower levels of sensitivity at the peak shoulders (green). (B) For complex mixtures, high abundance co-eluting peptide(s) (black), limits the time available for peptides of lower intensities (red). Peptide elution window and instrument cycling time (time necessary to switch between MS and MS/MS modes (as indicated by small purple boxes) also affect sensitivity.

With the recent development of MALDI instrumentation capable of both MS and MS/MS, significant efforts have been directed towards effectively coupling of LC to MALDI in an off-line manner. Although the addition of a fraction collection step can be viewed as a minor inconvenience, the off-line method offers a number of distinct advantages for the identification of low abundance proteins including peptides carrying post-translational modifications. With the decoupling of chromatographic and mass measurement steps, time constraints related to the LC elution window and instrument cycling time are no longer limiting factors. This feature of off-line LC-MALDI permits an entire LC run to be studied in MS mode prior to the acquisition of the first MS/MS spectra. While this protocol is not a replacement for high quality separations (as peptides still require detection), these features advantageously permit the targeted study of lower concentration peptides, regardless of the presence of higher abundance (co-eluting) analytes. Furthermore, because a semi-permanent representation of the LC run is generated, both hypothesis-generating and hypothesis-driven modes of research. This feature of the off-line MALDI protocol allows analyses to be conducted at higher levels of sensitivity because samples can be repeatedly re-examined without replicating sample injection(s). Furthermore, as precursor ions can now be selected based on any parameter including m/z , charge, ion abundance, retention time, or even after literature review, the technique offers an extremely flexible format to examine samples in “user-defined data-dependent” manner.

Despite the advantages of the off-line LC MALDI protocol, interfaces of this type are often in limited supply - presumably due to the high associated cost (\$40-60,000, Figure 1.22). As a result, although researchers may have institutional access to MALDI-MS/MS instrumentation, they may be restricted to PMF or MS/MS of high abundance peptides without the option of benefiting from the off-line LC MALDI protocol.

As an inexpensive alternative (\$3-5/sample) to LC-MALDI, the application of solid phase extraction (SPE) has been commonly applied to “clean” proteomic samples. Here, proteomic mixtures are passed through a reverse phase (C18) microcolumn housed within a pipette tip (ZipTip™, Millipore.com). Bound peptide material is then washed free of ionic components (detergents and salts to reduce MS interferences), and sample is eluted onto a MALDI target for study. While these methods are useful, SPE is incapable of resolving low abundance peptides from complex peptide mixtures, thus preventing their detection. Unfortunately, alternatives to SPE that are capable of true peptide separations are simply not available to proteomic researchers. To address this need, within Part 2 of this thesis, inexpensive, mechanically simple device for the off-line vacuum-driven chromatographic separation and deposition of peptides for MALDI-MS is described and demonstrated.

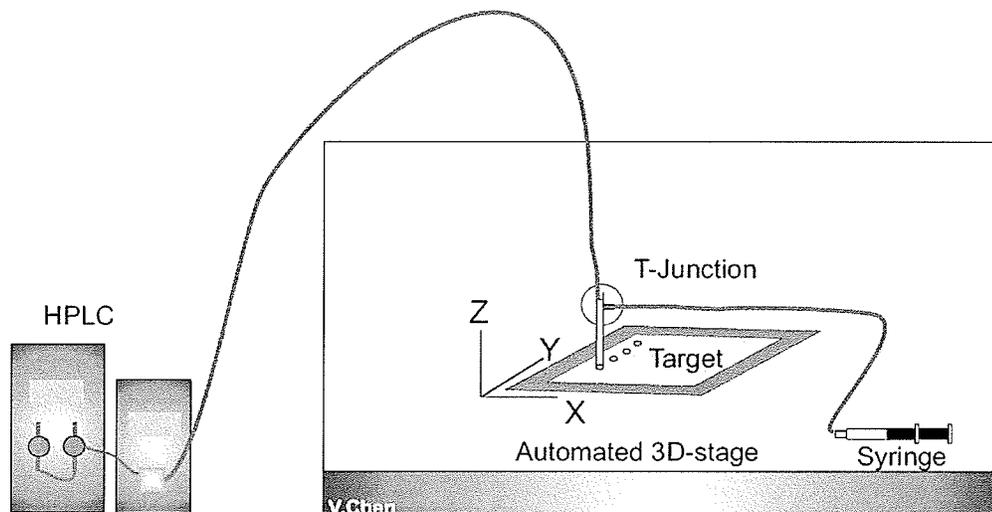


Figure 1.22: A typical off-line LC MALDI interface. A robotic stage capable of automated XYZ movement is programmed to collect and array HPLC effluent. To prepare samples for MALDI acquisitions, a syringe containing a matrix solution is mixed in-line with HPLC effluent. Samples are then allowed to dry prior to MS introduction.

The discovery of PDZ (please refer to section 1.2.3.1; Figure 1.3) containing protein zonula occludens-1 (ZO-1) at gap junctions and tight junctions with transcription factor ZONAB (Balda et al., 2000, Penes et al., 2005) and F-actin has recently expanded the role of these regions to include cytoskeletal regulation and transcriptional control (Fanning et al., 2002; Giepmans, 2004). Despite the functional importance of PDZ-domain proteins, bioanalytical methods of characterizing PDZ domains taking into account their biological selectivity for specific C-terminal sequences has not been

demonstrated. To address this issue, a portion of this work (Part 3) is focused on the characterization of the ZO-1 protein-protein interaction network, which included the use of the novel vacuum-driven LC deposition device (described in Part 2). Proteins isolated by affinity purification were identified and the physiological relevance of the newly discovered interactions were evaluated using a multifaceted approach which included MS, molecular-level analyses (IP, Co-IP, in-vitro pull-downs) and immunohistochemistry (Western-blot, colocalization analysis). Using these methods, a newly discovered interaction between ZO-1 and α -actinin-4 was identified and subsequently characterized. By identifying components of the ZO-1 scaffolding/signaling network, this work aims to help uncover possible modes of cell cycle regulation associated with the establishment of cell-cell and cell-matrix contacts.

Protein phosphorylation is a PTM that plays a key role in the regulation of cell signal transduction. Although phosphoproteomics has undergone rapid development, significant challenges still remain. For example, while proteomic studies employ hypothesis-generating workflows, researchers are increasingly called upon to identify/track phosphorylation in a hypothesis-driven manner. In other words, while proteomics is commonly used to identify unknown proteins within a sample, phosphoproteomic analyses are often conducted after the identities of proteins are established. Unfortunately, the automated MS/MS 'triggering' and data acquisition will often lead to a failure to detect low abundance phosphopeptides. While the off-line LC MALDI

protocol does offer some relief for the discovery of low abundance molecules, MS/MS triggering based on predicted phosphopeptide mass (peptide + 80Da) can still lead to a large number of candidate ions due to the complexity of proteomic mixtures. In other words, despite the advantages of off-line LC MALDI protocol, analyses often become laborious as validation of each candidate phosphopeptide masses can lead to unnecessary sample consumption (and corresponding losses in analytical sensitivity). Currently within the field of proteomics there are no targeted measures that can effectively address the targeted identification of phosphopeptides in LC-MS runs by either ESI- or MALDI-based techniques. To address these limitations, a method employing LC retention time and m/z was envisioned to provide criterion to reduce sample consumption in phosphoproteomic studies. Using standards mixtures, the Sequence Specific Retention Time Calculator (SSRCalc; Krokhin, 2006), was evaluated for its suitability to forecast phosphopeptide elution time. Upon completion of the proof-of-concept studies, phosphopeptides within *Drosophila* S2 cell lysates were analyzed, demonstrating the feasibility of the protocol with “real world” samples (Part 4).

1.5 References

- Aderem, A. (2005). Systems biology: its practice and challenges. *Cell*. 121, 511-3.
- Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature*. 422, 198-207.
- Alonso, A.; Sasin, J.; Bottini, N.; Friedberg, I.; Friedberg, I.; Osterman, A.; Godzik, a.; Hunter, T.; Dixon, J.; Mustelin, T. (2004). Protein tyrosine phosphatases in the human genome. *Cell*. 117, 699-711.
- Anderson, L.; Seilhamer, J. (1997). A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis*. 18, 533-537.
- Andersson L, Porath J. (1986) Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal Biochem*. 154, 250-4.
- Azarkan, M.; Huet, J.; Baeyens-Bolant, D.; Looze, Y.; Vandenbussche, G. (2007). Affinity Chromatography: a useful tool in proteomics studies. *J. Chromatogr. B Analyt Technol Biomed Life Sci*. 849, 81-90.
- Balda, M. S., Matter, K. (2000) The tight junction protein ZO-1 and an interacting transcription factor regulate ErbB-2 expression. *EMBO J*. 19, 2024-2033.
- Baldwin, M.A. (2005) Mass Spectrometers for the analysis of biomolecules. *Methods Enzymol*. 402, 3 - 48.
- Bateman, A.; Birney, E.; Cerruti, L. (2002). The Pfam protein families database. *Nucleic Acid Res*. 30; 276-80.
- Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J.E.; Villen, J.; Li, J.; Cohn, M.A.; Cantly, L.C.; Gygi, S.P. (2004). Large scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. USA*. 101, 12130-12135.
- Bezprozvanny, I.; Maximov, A. (2001). PDZ domains: More than just a glue. *Proc Natl Acad Sci USA*. 98, 787-789.
- Bhattacharyya, R.P.; Remenyi, A.; Yeh, B.J.; Lim, W.A. (2006). Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem*. 75, 655-80.
- Birrane, G.; Chung, J.; Ldias, J.A. (2003). Novel mode of ligand recognition by the Erbin PDZ domain. *J. Biol. Chem*. 278, 1399-1403.

- Blagoev, B. et al. (2003). A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nature Biotechnol.* 21, 315-318.
- Bonaventure, J.; Gibbs, L.; Home, W.C.; Baron, R. (2007). The localization of FGFR3 mutations causing thanatophoric dysplasia type I differentially affects phosphorylation, processing and ubiquitylation of the receptor. *FEBS J.* 274, 3078-90.
- Brown, R.S.; Lennon, J.J. (1995). Mass resolution improvement by incorporation of pulsed ion extraction in a linear time-of-flight mass spectrometer. *Anal. Chem.* 67, 1998-2003.
- Carr, S.; Aebersold, R.; Baldwin, M.; Burlingame, A.; Clauser, K.; Nesvizhskii, A.; Working group on publication guidelines for peptide and protein identification data (2004). The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. *Mol. Cell Proteomics.* 3, 531-3.
- Cao, P.; Stults, J.T. (1999). Phosphopeptide analysis by on-line immobilized metal-ion affinity chromatography-capillary electrophoresis-electrospray ionization mass spectrometry. *J. Chromatography.* 853, 225-235.
- Cekan, S.Z. (2004). Methods to find out the expression of activated genes. *Reprod. Biol. Endocrinol.* 2, 68-79.
- Celis, J.E.; Kruhoffer, M.; Gromova, I.; Frederiksen, C.; Ostergaard, M.; Thykjaer, T.; Gromov, P.; Yu, J.; Palsdottir, H.; Magnusson, N.; Orntoft, T.F. (2000). Gene expression profiling: Monitoring transcription and translation products using DNA microarrays and proteomics. *FEBS Lett.* 480, 2-16.
- Chaurand, P.; Luetzenkirchen, F.; Spengler, B. (1999). Peptide and protein identification by MALDI and MALDI-PSD Time-of-flight mass spectrometer. *J. Am. Soc. Mass Spectrom.* 10, 91-103.
- Chen, V.C.; Li, X.; Perreault, H.; Nagy, J.I. (2006). Interaction of zonula occludens-1 (ZO-1) with α -actinin-4 : Application of functional proteomic for the identification of PDZ domain-associated proteins. *J. Proteome Res.* 5, 2123-2134.
- Chernushevich, I.V.; Ens, W.; Standing, K.G. (1999). Orthogonal-injection time-of-flight mass spectrometry for the analysis of biomolecules. *Anal Chem.* 71., 452A-461A.
- Cho, K.O.; Hunt, C.A.; Kennedy, M.B. (1992). The rat brain postsynaptic density fraction contains a homolog of the *Drosophila* discs-large tumor suppressor protein. *Neuron.* 9, 929-942.

- Dawson, P.H. (ed.) (1995) *In* "Quadrupole mass spectrometry and its applications", American Institute of Physics, Woodbury, NY.
- Defilippi, P.; De Stefano, P.; Cabodi, S. (2006). p130Cas: a versatile scaffold in signaling networks. *Trends Cell Biol.* 16, 257-263.
- De Hoog, C.L.; Mann, M. (2004) Proteomics. *Annu. Rev. Genomics Hum. Genet.* 5, 267-293.
- Donaldson, L.W.; Gish, G.; Pawson, T.; Kay, L.E.; Forman-Kay, J.D. (2002). Structure of a regulatory complex involving the Abl SH3 domain, the Crk Sh2 domain and a Crk-derived phosphopeptide. *Proc. Natl. Acad. Sci. USA* 99, 14053-14058.
- Downard, K.M. (2006) Ions of the interactome: the role of MS in the study of protein interactions in proteomics and structural biology. *Proteomics.* 6, 5374-84.
- Doyle, D.A.; Lee, A.; Lewis, J.; Kim, E.; Sheng, M.; MacKinnon, R. (1996). Crystal structure of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell.* 85, 1067-76.
- Duckworth, H.E., Barber, R.C., Venkatasubramanian, Mass Spectroscopy, 2nd edn., Cambridge University Press, 1986, Chapter 5.
- El Rassi, Z.; Horvath, C. (1986) Metal chelate-interaction chromatography of proteins with iminodiacetic acid-bonded stationary phases on silica support. *J. Chromatogr.* 359, 241-53
- Ens, W.; Standing, K.G. (2005). Hybrid quadruple/time-of-flight mass spectrometers for analysis of biomolecules. *Methods Enzymol.* 402, 49-78.
- Everley, P.A.; Krijgsveld, J.; Zetter, B.R.; Gygi, S.P. (2004). Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol. Cell. Proteomics.* 3. 729-735.
- Everley, P.A.; et al. (2006) Enhanced analysis of metastatic prostate cancer using stable isotopes and high mass accuracy instrumentation. *J. Proteome Res.* 5, 1224-1231.
- Fanning, A. S., Ma, T. Y., Anderson, J. M. (2002) Isolation and functional characterization of the actin binding region in the tight junction protein ZO-1. *FASEB J.* 16, 1835-1837.

- Feller, S.M. (2001). Crk family adaptors-signalling complex formation and biological roles. *Oncogene*. 20, 6348-71.
- Feller, S.M.; Knudsen, B.; Hanafusa, H. (1994). C-Abl kinase regulated the protein binding activity of c-CRK. *Nature*. 13, 2341-2351.
- Fenn, J.B.; Mann, M.; Meng, C.K; Wong, S.F.; Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64-71.
- Ficarro, S.B.; McClelland, M.L.; Stukenberg, P.T.; Burke, D.J.; Ross, M.M.; Shabanowitz, J.; Hunt, D.F.; White, F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 20, 301-305.
- Futschik, M.E.; Chaurasia, G.; Herzel, H. (2007). Comparison of human protein-protein interaction maps. *Bioinformatics*. 23, 605-11.
- Gavin, A.C. et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*. 141-147.
- Gehrmann, M.L.; Hathout, Y.; Fenselau, C. (2004). Evaluation of metabolic labeling for comparative proteomics in breast cancer cells. *J. Proteome Res.* 3, 1063-1068.
- Gerhold, D.; Rushmore, T.; Caskey, C.T. (1999). DNA chips: Promising toys have become powerful tools. *Trend Biochem. Sci.* 24, 168-173.
- Giddings, J.C. (1984). Concepts and comparisons in multidimensional Separations. *Anal. Chem.* 56, 1258A.
- Giepmans, B.N. (2004). Gap junctions and connexins-interacting proteins. *Cardiovasc. Res.* 62, 233-45.
- Giepmans, B.N.; Moolenaar, W.H. (1998). The gap junction protein connexin43 interacts with the second PDZ domain of the zona occludens-1 protein. *Curr. Biol.* 8, 931-934.
- Gilar, M.; Olivova, P.; Daly, A.E.; Gebler, J.C. (2005) Orthogonality of separation in two-dimensional liquid chromatography. *Anal Chem.* 77, 6426-6434.
- Gimona, M. (2006). Protein linguistics – a grammar for modular protein assembly? *Nat. Rev. Mol. Cell Biol.* 7, 68-73.

- Gindhart, J.G. (2006). Towards and understanding of kinesin-1 dependent transport pathways through the study of protein-protein interactions. *Brief funct. genomic proteomic.* 5, 74-86.
- Gorg, A.; Weiss, W.; Dunn, M.J. (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics.* 4, 3665-85.
- Gronborg, M.; Kristiansen, T.Z., Stensballe, A.; Andersen, J.S.; Ohara, O.; Mann, M.; Jensen, O.N.; Pandey, A. (2002). A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phosphor-specific antibodies: identification of a novel protein, Frigg, as a protein kinase A substrate. *Mol. Cell Proteomics.* 1, 517-27.
- Guhaniyogi, J.; Brewer, G. (2001) Regulation of mRNA stability in mammalian cells. *Gene.* 265, 11-23.
- Guigo, R.; Flicek, P.; Abril, J.F.; Reymond, A.; Lagarde, J.; Denoeud, F.; Antonarakis, S.; Ashburner, M.; Bajic., V.B.; Birney, E.; Castelo, R.; Eyraes, E.; Ucla, C.; Gingeras, T.R.; Harrow, J.; Hubbard, T.; Lewis, S.E.; Reese, M.G. (2006). EGASP: the human ENCODE Genome Annotation Assessment Project. *Genome Biology.* 7, S2.
- Gygi, S.P.; Corthals, G.L.; Zhang, Y.; Rochon, Y.; Aebersold, R. (2000). Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9390-9395.
- Hegde, P.S.; White, I.R.; Debouck, C. (2003) Interplay of transcriptomics and proteomics. *Curr. Opin. Biotechnol.* 647-651.
- Ho, Y. et al. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature.* 415, 180-183.
- Imanishi, T.; Itoh, T.; Suzuki, Y.; O'Donovan, C.; Fukuchi, S.; Koyanagi, K.O.; Barrero, R.A.; Tamura, T.; Yamaguchi-Kabata, Y.; Tanino, M. et al. (2004) Integrative annotation of 21,037 human genes validated by full-length cDNA clones. *PLoS Biol.* 2, 856-875.
- Issaq, H.J.; Chan, K.C.; Janini, G.M.; Conrads, T.P.; Veenstra, T.D. (2005). Multidimensional separation of peptides for effective proteomic analysis. *J. Chromatogr. B.* 817, 35.
- Itoh, M., Furuse, M., Morita, K., Kubota, K., Saitou, M., Tsukita, S. (1999) Direct binding of three tight junction-associated MAGUKs, ZO-1, ZO-2, and ZO-3, with the COOH termini of claudins. *J. Cell Biol.* 147, 1351-1363.

- Jiang, W.; Hern, M.T.W. (1996). Protein interaction with immobilized metal ion affinity ligands under high ionic strength conditions. *Anal. Biochem.* 242, 45-54.
- Johnson, R.S.; Martin, S.A.; Biemanann, K.; Stults, J.I.; Watson, J.T. (1987). Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* 59, 2621-5.
- Jones RB, Gordus A, Krall JA, MacBeath G. (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature.* 439, 168–74.
- Karas, M.; Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 Daltons. *Anal. Chem.* 60, 2299-301.
- Karty, J.A.; Ireland, M.M.; Brun, Y.V.; Reilly, J.P. (2002). Artifacts and unassigned masses encountered in peptide mass mapping. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 782, 363-383.
- Kaufmann, R.; Kirch, D.; Spengler, B. (1994). Sequencing of peptides in a time-of-flight mass spectrometer: Evaluation of postsource decay following matrix-assisted laser desorption ionization (MALDI). *Int. J. Mass Spectrom. Ion Proc.* 131, 355-385.
- Kausalya, P.J.; Reichert, M.; Hunziker, W. (2001). Connexin45 directly binds to ZO-1 and localizes to the tight junction region in epithelial MDCK cells. *FEBS Lett.* 505, 92-96.
- Kim, E.; Niethammer, M.; Rothschild, A.; Jan, Y.N.; Sheng, M. (1995). Clustering of Shaker-type K⁺ channels by interaction with a family of membrane-associated guanylate kinase. *Nature.* 378, 85-88.
- Klose, J., Kobalz, U. (1995). Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis.* 16, 1034-59.
- Kornau, H.-C.; Schenker, L. T.; Kennedy, M. B.; and Seeburg, P. H. (1995). Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95. *Science.* 269, 1737-1740.
- Krishna, R.G.; Wold, F. (1998). Posttranslational modifications. *Proteins – Analysis and design.* R. H. Angeletti. San Diego, Academic Press. 2, 121-206.

- Krokhin, O.V.; Ens, W.; Standing, K.G. (2005). MALDI QqTOF MS combined with off-line HPLC for characterization of protein primary structure and post-translational modifications. *J. Biomol Tech.* 16, 459-440.
- Krokhin, O.V. (2006) Sequence-specific retention time Calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: Application to 300- 100A Pore Size C18 Sorbents. *Anal. Chem.* 78 7785-7795.
- Krutchinsky, A.N.; Loboda, A.V.; Spicer, V.L.; Dworschak; R.; Ens, W.; Standing, K.G.; (1998)^a. Orthogonal injection of MALDI ions into a time-of-flight spectrometer through a collisional damping interface. *Rapid Commun. Mass Spectrom.* 12, 508-518.
- Krutchinsky, A.N.; Chernushevich, I.V.; Spicer, V.; Ens, W.; Standing, K.G. (1998)^b. A collisional damping interface for an electrospray ionization TOF mass spectrometer. *J. Am. Soc. Mass Spectrom.* 9, 569-579.
- Larsen, M.R.; Sorensen, G.L.; Fey, S.J.; Larsen, P.M., and Roepstorff, P. (2001). Phospho-proteomics: Evaluation of the use of enzymatic de-phosphorylation and differential mass spectrometric peptide mass mapping for site specific phosphorylation assignment in proteins separated by gel electrophoresis. *Proteomics.* 1, 223-238.
- Larsen, M.R.; Thingholm, T.E.; Jensen, O.N.; Roepstorff, P.; Jorgensen, T.J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell Proteomics.* 4, 873-86.
- Li, X.; Olson, C.; Lu, S.; Kamasawa, N.; Yasumura, T.; Rash, J.E.; Nagy, J.I. (2004). Neuronal connexin36 association with zonula occludens-1 protein (ZO-1) in mouse brain and interaction with the first PDZ domain of ZO-1. *Eur J Neurosci.* 19, 2132-46.
- Link, A.J.; Eng, J.; Schieltz, D.M.; Carmack, E.; Mize, G.J.; Morris, D.R.; Garvik, B.M.; Yates, J.R. (1999). Direct analysis of protein complexes using mass spectrometry. *Nature biotechnol.* 17, 676-682.
- Luo, Q., Nieves, E., Kzhyskowska, J., Angeletti, R. H. (2006). Endogenous transforming growth factor-beta receptor-mediated Smad signaling complexes analyzed by mass spectrometry. *Mol. Cell. Proteomics.* 5, 1245.
- Mamyrin, B.A.; Shmikk, D.V. (1979). The linear mass reflectron. *Zh. Eksp. Teor. Fiz.* 76, 1500.

- Mann, M.; Hojrup, P.; Roepstorff, P. (1993). Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol. Mass Spectrom.* 22, 338-45.
- Matter, K., Balda, M.S. (2003). Signalling to and from tight junctions. *Nat. Rev. Mol. Cell. Biol.* 4, 225-236.
- McLachlin, D.T.; Chait, B.T. Analysis of phosphorylated proteins and peptides by mass spectrometry. *Curr. Opin. Chem. Biol.* 5, 591-602.
- Medzihradzky, K.F.; Campbell, J.M.; Baldwin, M.A.; Falick, A.M.; Juhasz, P.; Vestal, M.L. Burlingame, A. (2000). The characteristics of peptide collision induced dissociation using a high-performance MALDI TOF/TOF tandem mass spectrometer. *Anal. Chem.* 72, 552-558.
- Mitic, L. L., Schneeberger, E. E., Fanning, A. S., Anderson, J. M. (1999) Connexin-occludin chimeras containing the ZO-binding domain of occludin localize at MDCK tight junctions and NRK cell contacts. *J. Cell Biol.* 146, 683-693.
- Morais-Cabral, J.H.; Petosa, C.; Sutcliffe, M.J.; Raza, S.; Byron, O.; Poy, F.; Marfatia, S.M.; Chishti, A.H.; Liddington, R.C. (1996). Crystal structure of a PDZ domain. *Nature.* 382, 649-652.
- Mukherji, M.; Brill, L.M.; Ficarro, S.B.; Hampton, G.M.; Schultz, P.G. (2006) A Phosphoproteomic Analysis of the ErbB2 Receptor Tyrosine Kinase Signaling Pathways. *Biochemistry* 45, 15529-15540.
- Neubauer, G.; King, A.; Rappsilber, J.; Calvio, C.; Watson, M.; Ajuh, P.; Sleeman, J.; Lamond, A.; Mann, M. (1998). Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat. Genet.* 20, 46-50.
- Neverova, I., van Eyk, J.E. (2005). Role of Chromatographic techniques in proteomic analysis. *J. Chromatogr. B.* 815, 51-63.
- Nordhoff, E.; Krogsdam, A.M.; Jorgensen H.F.; Kallipolitis, B.H.; Clark, B.F.; Roepstorff, P.; Kristiansen, K. (1999) Rapid identification of Dana-binding proteins by mass spectrometry. *Nat. Biotechnol.* 17, 884-888.
- Nuwaysir, L.M.; Stults, J.T. (1993). Electrospray ionization mass spectrometry of phosphopeptides isolated by on-line immobilized metal-ion affinity chromatography. *J. Am. Soc. Mass Spectrom.* 4, 662-669.
- O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007-21.

- Ong, S.E. et al. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* 1, 376-386.
- Ong, S.E.; Foster, L.J.; Mann, M. (2003)^a. Mass spectrometric-based approaches in quantitative proteomics. *Methods*. 29, 124-130.
- Ong, S.E.; Kratchmarova, I.; Mann, M. (2004)^bProperties of ¹³C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* 2, 173-181.
- Ong, S.E.; Mann, M. (2005). Mass spectrometry-based proteomics turns quantitative. *Nature Chem. Boil.* 1. 252-262.
- Paizs, B.; Suhai, S. (2005). Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* 24, 508-548.
- Pandey, A.; Fernandez, M.M.; Steen, H.; Blagoev, B.; Nielsen, M.M.; Roche, S.; Mann, M.; Lodish, H.F. (2000)^a. Identification of a novel immunoreceptor tyrosine-based activation motif-containing molecule, STAM2, by mass spectrometry and its involvement in growth factor and cytokine receptor signaling pathways. *J. Biol. Chem.* 275, 38633-9.
- Pandey, A.; Mann, M. (2000)^b. Proteomics to study genes and genomes. *Nature*. 405, 837-46.
- Pandey, A.; Podtejnikov, A.V.; Blagoev, B.; Bustelo, X.R.; Mann, M.; Lodish, H.F. (2000)^c. *Proc. Natl. Acad. Sci. USA.* 97, 179.
- Patton, W.F. (2002). Detection technologies in proteome analysis. *J. Chromatogra. B Analyt. Technol. Biomed. Life Sci.* 771, 3-31.
- Pawson, T. (2004). Specificity in Signal Transduction: From Phosphotyrosine-SH2 Domain Interactions to Complex Cellular Systems. *Cell*. 116, 191-203.
- Pawson, T. (2007). Dynamic control of signaling by modular adaptor proteins. *Curr Opin Cell Biol.* 19, 112-116.
- Pawson, T.; Nash, P. (2000). Protein-protein interactions define specificity in signal transduction. *Genes Dev.* 14, 1027-1047.
- Peng, J.; Gygi, S.P. (2001). Proteomics: the move to mixtures. *J. Mass Spectrom.* 36, 1083-1091.

- Perkins, D.N.; Pappin, D.J.; Creasy, D.M.; Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 20, 3551-3556.
- Peters, B.A.; St. Croix, B.; Sjoblom, T.; Cummins, J.M.; Silliman, N.; Ptak, J.; Saha, S., Kinzler, K.W.; Hatzis, C.; Velculescu, V.E. (2007) Large-scale identification of novel transcripts in the human genome. *Genome Res.* 17, 287-292.
- Penes, M., Li, X., Nagy, J.I. (2005). Expression of zonula occludens-1 (ZO-1) and transcription factor ZO-1-associated nucleic acid-binding protein (ZONAB)-MsY3 in glial cells and colocalization at oligodendrocyte and astrocyte gap junctions in mouse brain. *Euro. J. Neurosci.* 22, 404-418.
- Pennisi, E. (2007). Working the (gene count) numbers: Finally a firm answer? *Science.* 316, 1113.
- Picotti, P., Aebersold, R., Domon, B. (2007). The Implications of Proteolytic Background for Shotgun Proteomics. *Mol. Cell. Proteomics. MCP-online.*
- Plumb, R.S.; Rainville, P.; Smith, B.W.; Johnson, K.A.; Castro-Perez, J.; Wilson, I.D.; and Nicholson, J.K. (2006) Generation of ultrahigh peak capacity LC separations via elevated temperatures and high linear mobile-phase velocities. *Anal. Chem.* 78, 7278-83 .
- Ponting, C.P. (1997). Evidence for PDZ domains in bacteria, yeast and plants. *Protein Sci.* 6, 464-468.
- Posewitz, M.C.; Tempest, P. (1999). Immobilized gallium (III) affinity chromatography of phosphopeptides. *Anal. Chem.* 71, 2883-2892.
- Raska, C.S.; Parker, C.E.; Dominski, Z.; Marzluff, W.F.; Glish, G.L.; Pope, R.M.; Borchers, C.H. (2002). Direct MALDI-MS/MS of phosphopeptides affinity-bound to immobilized metal ion affinity chromatography. *Anal. Chem.* 74, 3429-33.
- Reinders, J.; Sickmann, A. (2005). State-of-the-art in phosphoproteomics. *Proteomics.* 5, 4052-61.
- Rigaut, G. et al. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnol.* 17, 1030-1032.
- Robertson, S.C.; Tynan, J.A; Donoghue, D.J. (2000). RTK mutations and human syndromes: When good receptors turn bad. *Trends Genet.* 16, 265-71.

- Rosen, M.K.; Yamazaki, T.; Gish, G.D.; Kay, C.M.; Pawson, T.; Kay, L.E. (1995). Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the signaling protein CRK. *Nature*. 374, 477-479.
- Rout, M.P.; Aitchison J.D.; Suprpto, A.; Hjertaas, K.; Zhao, Y.; Chait, B.T. (2000). The yeast nuclear pore complex: Composition, architecture, and transport mechanism. *J. Cell Biol.* 148, 635-652.
- Roepstorff, P.; Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* 11, 601.
- Ross, J. (1995). mRNA stability in Mammalian cells. *Microbiol Rev.* 59, 423-450.
- Sawada, Y.; Tamada, M. Bublin-Thaler, B.J.; Cherniavskaya, O.; Sakai, R.; Tanaka, S.; Sheetz, M.P. (2006). Force sensing by mechanical extension of the Src family kinase substrate p130Cas. *Cell* 127, 1015-1026.
- Scherperel, G.; Reid, G.E. (2007). Emerging methods in proteomics: top-down protein characterization by multistage tandem mass spectrometry. *Analyst*. 132, 500-6.
- Schultz, J.; Copley, R.R.; Doerks, T.; Ponting, C.P.; Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* 28, 231-234.
- Schulze, W.X.; Mann, M. (2004). A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* 279, 10756-10764.
- Schulze, W.X.; Deng, L.; Mann, M. (2005). Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol. Syst. Biol.* 1, E1-E13.
- Skeet, B.T.; Dikic, I.; Zhou, M.-M.; Pawson, T. (2006) Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* 7, 473-83.
- Smith, D.B.; Johnson, K.S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67, 31-40.
- Songyang, Z.; Fanning, A.S.; Fu, C.; Xu, J.; Marfatia, S.M.; Chishti, A.H.; Crompton, A.; Chan, A.C. Anderson, J.M.; Cantley, L.C. (1997). Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science*. 275, 73-77.

- Spengler, B.; Cotter, R.J. (1990). Ultraviolet laser desorption/ionization mass spectrometry of protein above 100,000 Daltons by pulsed ion extraction time-of-flight analysis. *Anal. Chem.* 62, 793-796.
- Spengler, B.; Kirsch, D.; Kaufmann, R.; Jaeger, E. (1992). Peptide sequencing by matrix-assisted laser desorption mass spectrometry. *Rapid Commun. Mass Spectrom.* 6, 105-108.
- Stannard, C.; Soskic, B.; Godovac-Zimmermann, J. (2003). Rapid changes in the phosphoproteome show diverse cellular responses following stimulation of human lung fibroblasts with endothelin-1. *Biochemistry.* 42, 13909-18.
- Staudinger, J.; Lu, J.; Olson, E.N. (1997) Specific interaction of the PDZ domain protein PICK1 with the COOH terminus of protein kinase C-alpha. *J. Biol. Chem.* 272, 32019-24.
- Steen, H.; Kuster, B.; Fernandez, M.; Pandey, A.; Mann, M. (2002). Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway. *J. Biol. Chem.* 277, 1031-9.
- Stricker, N. L.; Christopherson, K. S.; Yi, B. A.; Schatz, P. J.; Raab, R. W.; Dawes, G.; Bassett, D. E., Jr.; Bredt, D. S.; Li, M. (1997) PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences. *Nat. Biotechnol.* 15, 336-342.
- Stults J.T.; Arnott, D. (2005) Proteomics. *Methods Enzymol.* 402, 245-89.
- Takino, T.; Tamura, M.; Miyamori, H.; Araki, M.; Matsumoto, K.; Sato, H.; Yamada, K.M. (2003). Tyrosin phosphorylation of the CrklI adaptor protein modulates cell migration. *J. Cell Sci.* 116, 3145-3155
- Tannu, N.S.; Hemby, S.E. (2006) Two-dimensional fluorescence difference gel electrophoresis for comparative proteomics profiling. *Nat. Protoc.* 1, 1732-42.
- Thiede, B.; Hohenwarter, W.; Krah, A.; Mattow, J.; Schmade, M.; Schmidt, F.; Jungblut, P.R. (2005). Peptide mass fingerprinting. *Methods.* 35, 237-247.
- Thomson, J. J. Rays of Positive Electricity and Their Application to Chemical Analyses, Green & Co., London, 1913.
- Thomson, J.J. Rays of positive electricity and their application to chemical analyses, Green & Co., London, 1923.
- Tochio, H.; Zhang, Q.; Mandal, P.; Li, M.; Zhang, M. (1999) Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide. *Nat. Struct. Biol.* 6, 417-421.

- Tu, L.C.; Yan, X.; Hood, L.; Lin, B. (2007) Proteomics analysis of the interactome of N-myc downstream regulated gene 1 and its interactions with the androgen response program in prostate cancer cells. *Mol. Cell Proteomics*. 6, 575-88.
- Vanrobaeys, F.; Van Coster, R.; Dhondt, G.; Devreese, B.; Van Beeumen, J. (2005). Profiling of myelin proteins by 2D-gel electrophoresis and multidimensional liquid chromatography coupled to MALDI TOF-TOF mass spectrometry. *J. Proteome Res.* 4, 2283-93.
- Van Zeijl, L.; Ponsioen, B.; Giepmans, B.N.G.; Ariaens, A.; Postma, F.R.; Varnai, P.; Ballen, T.; Divecha, N.; Jalink, K.; Moolenaar, W.H. (2007). Regulation of connexins43 gap junctional communication by phosphatidylinositol 4,5-bisphosphate. *J. Cell Biol.* 177, 881-891.
- Venter, J.C.; Adams, M.D., Myers, E.W.; Li, P.W., Mural, R.J. (2001) The sequence of the human genome. *Science*. 291, 1304-1351.
- Vestal, M.L.; Campbell, J.M. (2005) Tandem time-of-flight mass spectrometry. *Methods Enzymol.* 402, 79 - 108.
- Vestal, M.L.; Juhasz, P; Martin, S.A. (1995). Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 9, 1044-1050.
- Von Mering, C. et al. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. *Nature*. 417, 399-403.
- Wang, H.; Hanash, S. (2003) Multi-dimensional liquid phase based separation in proteomics. *J. Chromatogr. B.* 787, 11-18.
- Washburn, M.P., Wolters, D. and Yates, J.R. 3 (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19, 242-7.
- Wiley, W.C.; McLaren, I.H. (1955). Time-of-flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* 26, 1150.
- Wilkins, M., Sanchez, J., Gooley, A., Appel, R., Humphry-Smith, I., Hochstrasser, D., and Williams, K. (1995). Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* 13, 19-50.

- Wittmann-Liebold B., Graack, H.R., Pohl, T. (2006) Two-dimensional gel electrophoresis as tool for proteomics studies in combination with protein identification by mass spectrometry. *Proteomics*. 6, 4688-703.
- Wolters, D.A.; Washburn, M.P.; Yates, J.R. 3rd (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* 73, 5683-90.
- Zhang, H.; Yan, W.; Aebersold, R. (2004) Chemical probes and tandem mass spectrometry: a strategy for the quantitative analysis of proteomes and subproteomes. *Curr. Opin. Chem. Biol.* 8, 66-75.
- Zhang, M.; Wang, W. (2003). Organization of Signaling complexes by PDZ-Scaffolding Proteins. *Acc. Chem. Res.* 36, 530-538.
- Zheng, H.; Ju, P.; Quinn, D.F.; Wang, Y.K. (2005). Phosphotyrosine proteomic study of interferon alpha signaling pathway using a combination of immunoprecipitation and immobilized metal affinity chromatography. *Mol. Cell Proteomics*. 4, 721-730.
- Zhu, W.; Reich, C.I.; Olsen, G.J.; Giometti, C.S.; and Yates, J.R.; 3rd. (2004). Shotgun proteomics of methanococcus jannaschii and insights into methanogenesis. *J. Proteome Res.* 3, 538-548.
- Zonlnierowicz, S.; Bollen, M. (2000) Protein phosphorylation and protein phosphatases. *EMBO J.* 19, 483-8.

Part 2:

Development of an off-line vacuum-LC/deposition device for MALDI-MS and
MS/MS

2.0 Author's Contributions

I designed and conducted all experiments within this section including patented materials found in Part 6. Operation of the MALDI-QqTOF in MS/MS mode was provided by K. Cheng of the Manitoba Center for Proteomics Time-of-flight lab directed by K.G. Standing and W. Ens. I was also responsible for the figures and writing manuscript drafts. The final version was edited by H. Perreault prior to submission to the journal Analytical Chemistry.

-V. Chen

A Device for the Reverse Phase Separation and on-target Deposition of Peptides Incorporating a Novel Hydrophobic Sample Barrier for Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry*

Vincent C. Chen¹, Keding Cheng², Werner Ens³, Kenneth G. Standing³, James I. Nagy⁴ and H  l  ne Perreault¹

1. Department of Chemistry, University of Manitoba
Winnipeg, Manitoba, Canada

2. Manitoba Center for Proteomics, University of Manitoba
Winnipeg, Manitoba, Canada

3. Department of Physics and Astronomy, University of Manitoba
Winnipeg, Manitoba, Canada

4. Department of Physiology, University of Manitoba
Winnipeg, Manitoba, Canada

*Published: **Analytical Chemistry** 2004 February 15; 76 (4): 1189-96*

Portions reproduced with permission from **Analytical Chemistry 2004 February 15; 76 (4): 1189-96.*

Copyright   2004 American Chemical Society

2.1 Abstract

The separation of peptide mixtures from proteolytic cleavage is often necessary prior to mass spectrometry (MS) to enhance sensitivity and peptide mapping coverage. When buffers, salts and other higher abundance peptides/contaminants are present, competition for charge during the electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) processes can lead to ion suppression for the targeted analyte(s). In this article, a simple reverse phase micro-column sample separation and deposition device (SepDep) is described. The use of this device improves or renders possible the analysis of complex or contaminated peptide mixtures by MALDI-MS. The method is simple, inexpensive and utilizes single-use low-cost Geloader™ type columns packed with reverse phase material. The device described utilizes an open column, allowing for a gradient or narrow-step gradient to be applied by any solvent delivery system, or manually with a pipette. A key feature of the device is a deposition chamber that can be custom-built to hold any MALDI target. The SepDep is attached directly to an in-house vacuum line and draws solvent from the open-ended LC column. The elution of separated peptides is performed directly onto a target that has been treated with a hydrophobic barrier. This barrier effectively isolates fractions and improves the quality and morphology of the matrix crystals. The method produces efficient separations of proteolytic peptides, significantly reducing signal suppression effects in MALDI.

2.2 Introduction

Mass spectrometry (MS) has become a major analytical tool in proteomic and biological research. Most protein identification strategies involving MS analyze proteolytic peptides (e.g. tryptic digests) for mass fingerprinting in combination with tandem MS (MS/MS) to confirm amino acid sequence and the presence of various post-translational modifications. In most proteome studies, proteins are separated on electrophoretic gels and in-gel digestion extracts are subjected to MS analysis. Although matrix-assisted laser desorption/ionization (MALDI) is very effective for screening high-abundance proteins in complex samples, lower abundance peptides often remain undetected (Ho et al., 2002). Suppression effects are a common problem, arising from the presence of multiple analytes competing for protons during the ionization process. These effects are mostly pronounced when the abundance of analytes of interest decreases relative to more concentrated peptides and/or contaminants present in the sample matrix (Krause et al., 1999; Westman et al., 1997; Zhang et al., 1999; Gygi et al., 2000; Wall et al., 2002). Autolytic cleavage of the protease and presence of peptides from human keratins also constitute potential problems.

Wilm and Mann (1994) were the first to use micro-columns filled with reverse phase media from solid phase extraction (SPE) and direct elution/deposition into electrospray ionization (ESI) needles. Gobom et al. (1999) used Geloader™ type C18 columns to develop a method capable of

concentrating femto- to attomoles of peptides using packing beds of 30-50 nL. These methods have gained wide acceptance and recognition for their simplicity and efficiency for removing ionic salts, buffers and detergents, thus enhancing the quality of MS analyses.

Although SPE methods allow for the removal of ionic and polar components from the sample matrix, suppression effects resulting from highly abundant peptides are often problematic. The batch elution of peptide extracts using increasing percentages of organic solvent (sequential elution with e.g. 25, 50, and 75% of acetonitrile in water) helps to limit suppression, but in the case of complex samples, a significant number of peptides may elute over a given solvent batch range. The use of narrow solvent percentage increments may solve these problems, but can be time consuming and not well adapted to direct on-target deposition for MALDI-MS analysis.

The removal of interfering compounds from peptide mixtures by continuous flow liquid chromatography (LC) or capillary electrophoresis (CE) is common practice to minimize ion suppression. The separation of peptide mixtures by LC or CE also allows the resolution of isobaric species and isoforms that otherwise may be difficult to detect or distinguish. LC separation prior to MS improves sequence coverage (Apffel et al., 1995; Grimm et al., 1998; Heieh et al., 1998; Udiavar et al., 1998; Mauri et al., 1999; Chong et al., 2001; Bratt et al., 2001) and enhances the quality of the analysis (Preisler et

al., 2000; Foret et al., 2002) while allowing the identification of larger peptide fragments that were not detectable in the presence of smaller ones (Reid et al., 1998; Rusconi et al., 1998).

In this work, the advantages of coupling LC separations with MALDI have been exploited in the design and development of a simple and inexpensive device for the collection of peptide fractions. This device serves the function of a micro-LC system by separating peptides using reverse phase material packed in a Geloader™ type column (Chen et al., 2003a; Chen et al., 2003b). The design of this device includes a clear acrylic vacuum chamber to house the MALDI target.

The chamber is an open bottom (5-sided) box placed onto a smooth polyvinyl chloride (PVC) sheet. The box has a connection to an in-house vacuum line with rubber tubing. The center of the top panel has a small hole that securely fastens the column. The top panel also has a 'thumb controlled' release valve used to position the deposition device, to control the flow rate, and to facilitate target removal. It is important to note that the MALDI target and PVC sheet are stationary to the bench-top and that the clear acrylic box (including column) is positioned for sample deposition. When the chamber is evacuated, solvent is drawn through the column and peptides are eluted directly onto the target located within the SepDep device. Features of the device allow the use of a narrow-step gradient with a pipette. The volume

deposited per spot can be controlled by changing the initial distance (when the chamber is at atmosphere) between the end of the column and the target, and/or by controlling vacuum levels within the chamber using the thumb valve.

Sample deposition for MALDI is aided by the pre-application of a hydrophobic barrier with a commercially available pen. This highly hydrophobic barrier is composed of polybutadiene (butadiene rubber) and is easily applied onto the target using a marker-type applicator for fraction partitioning. The barrier is reusable and remains highly hydrophobic for several application/wash cycles. It also increases the uniformity of the peptide dispersed within the MALDI matrix, reducing edge effects. Using the barrier, DHB matrix can be pre-applied and dried onto a 384 spot target in less than 4 minutes prior to the deposition of peptide samples. The typical volume-per-spot deposited on the target was between 1.0 and 1.5 μL . Collection of 200-344 fractions was completed within 30-45 minutes using a 0-80% acetonitrile:H₂O (0.1% TFA) narrow-step gradient for MS and MS/MS. The cost of construction for this SepDep device, not including disposables, was under \$70.00 US. All materials were readily available from local hardware and scientific suppliers.

For the purposes of this article, the SepDep device was tested first to demonstrate its utility in reducing MALDI ion suppression effects. The device

was then used to map tryptic peptides from citrate synthase and phosphorylated β -casein.

2.3 Experimental

2.3.1 Materials.

Geloader™ tips were purchased from Eppendorf (Hamburg, Germany). Trifluoroacetic acid (TFA), phosphorylated bovine β -casein, L-1-tosylamide-2-phenylethylchloromethyl ketone (TPCK) treated bovine pancreatic trypsin, 2,5-dihydroxybenzoic acid (DHB) and ammonium bicarbonate were purchased from Sigma Chemicals (St. Louis, MO). Connexin-43 peptide corresponding to the sequence CDQRPSSRASSRASSRPR was synthesized by Anaspec (San Jose, CA). The 1 mL syringes used to fabricate O-rings were obtained from Becton Dickinson (Franklin Lakes, NJ). Fluoropore™ PTFE membrane filter disks with 0.45 μ m pore size were obtained from Millipore (Bedford, MA). Widepore™ (15 μ m, 300 Å pore size) C18 packing material used in these experiments was purchased from JT Baker (Phillipsburg, NJ). HPLC grade acetonitrile (ACN) and methanol (MeOH) were obtained from Fisher Scientific (Fair Lawn, NJ). Deionized, distilled water was prepared with a Barnstead Nanopure (Dubuque, IA) water filtration/deionization system supplied by a reversed osmosis feed stock. All solutions were degassed by sonication under vacuum for at least 10 minutes.

2.3.2 Peptides used for ion suppression study.

HPLC purified Connexin-43 peptide (Anaspec, 2 µg, 1 nmol) was dissolved in 10 µL of 20 mM aqueous ammonium bicarbonate solution. The solution was digested by the addition of 0.5 µL of trypsin (1.0 mg/mL) for 1 hour at 37°C. Digestion was halted by the addition of TFA to a final concentration of 2.5%. The solution was then spiked with 5.0 µL of angiotensin I in water (3.85×10^{-3} M) and mixed thoroughly. The solution (2.0 µL) was then transferred to the column for separation and study.

2.3.3 Digestion of β-casein.

Phosphorylated bovine β-casein (2.5 mg, 10 nmol) was dissolved in 500 µL of aqueous 20 mM ammonium bicarbonate. TPCK treated trypsin (1.0 mg/mL) in 20 mM ammonium bicarbonate was then added to the solution, to a final relative mass ratio of 50:1 (β-casein:trypsin) and the solution was allowed to incubate for 14 hours at 37°C. Digestion was stopped by adding TFA to a final concentration of 2.5% and stored at -20°C until application to the reverse phase column. Once the RP column was prepared, 2.5 µL of the β-casein digest (2.5 nmol) was loaded onto the column and separated as described later.

2.3.4 Digestion of *E. coli* citrate synthase.

A stock solution of citrate synthase (257 mg/mL) in 20 mM Tris-HCl, 1 mM EDTA and 50 mM KCl was prepared as described by Ayed *et al.* (1998).

Citrate synthase stock solution (3 μL) was then diluted to 50 μL with 20 mM ammonium bicarbonate to a concentration of 3.2×10^{-4} M. The protein was digested by adding 15 μL of 1mg/mL trypsin and incubating for 4 hours at 37°C. Digestion was halted by adding 3 μL of TFA and the solution was stored at -20°C. The peptide solution (4.0 μL) was then loaded onto the RP column for on-target deposition of separated peptides using the SepDep device. Resulting fractions were then analyzed by MALDI-MS and resulting peptides masses were analyzed.

2.3.5 SepDep device.

The chamber was constructed out of clear acrylic plastic. The open bottom acrylic box measures L25.0 cm x W25.0 cm x H4.5 cm, is held together with screws, and is sealed with general-purpose household silicone. The open edge of the box is lined with closed cell weather stripping (0.6 x 1.0 cm) to provide a good seal when placed upon a smooth polyvinyl chloride sheet (PVC, 54 cm x 30 cm, Figure 2.1). Both acrylic and PVC sheets are 1.2 cm thick and provide sufficient rigidity when placed under maximum house vacuum (0.2 atm, University of Manitoba, Department of Chemistry). The column is inserted into the center-top surface and sealed with an O-ring constructed from that of a 1 mL syringe. Vacuum within the chamber is established using a screw-in tapered hose connector and standard quality rubber tubing. A screw-in stainless steel tube (thumb valve) has been installed near the corner of the top surface of the device.

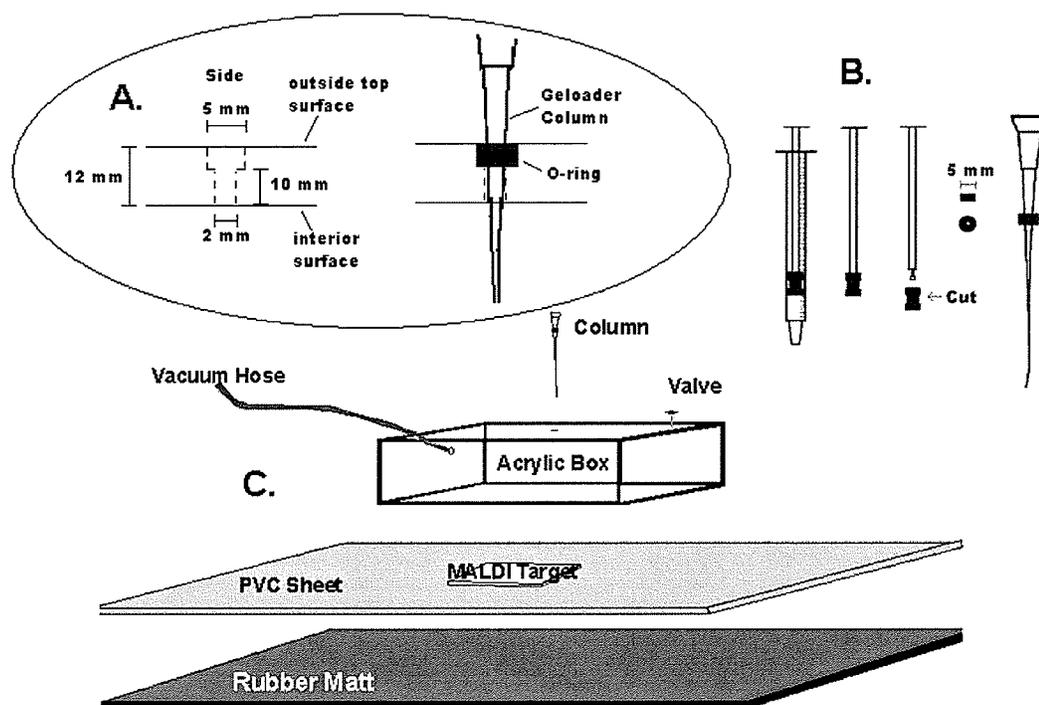


Figure 2.1: Components of the SepDep device. A) The column is fixed to the device by drilling a 2 mm diameter hole through the center-top panel of the deposition device. A larger 5 mm diameter tap is centered onto the 2 mm hole. B) A standard plastic 1 mL syringe is dismantled and the plunger is removed. The plunger is cut at the position indicated to form an O-ring. The Geloader™ column is inserted through the O-ring (approximately 2.0 – 2.75 cm from the top of the column) and positioning is used to customize column distance from the target. The modified column is inserted into the center top hole of the device, creating a tight seal. C) Components of the SepDep device used in these experiments. A rubber mat is used to prevent slippage of the PVC sheet on the bench top.

2.3.6 Hydrophobic MALDI target preparation.

The hydrophobic barrier was applied to a stainless steel Bruker Daltonics 384 Scout™ target (Figure 2.2). The Super HT™ Pap Pen, with a 4 mm tapered point, was manufactured by Daido Sangyo Co. Ltd (Tokyo, Japan). A 4 mm wide barrier was ‘drawn’ onto the target using parallel lines to separate the rows (A-P) and columns (1-24). The target was then rinsed with approximately 200 mL of water followed by 200 mL of methanol. The new surface was then allowed to cure at room temperature for at least 24 hours under vacuum. The target was then stored in a dust-free environment to prevent surface contamination.

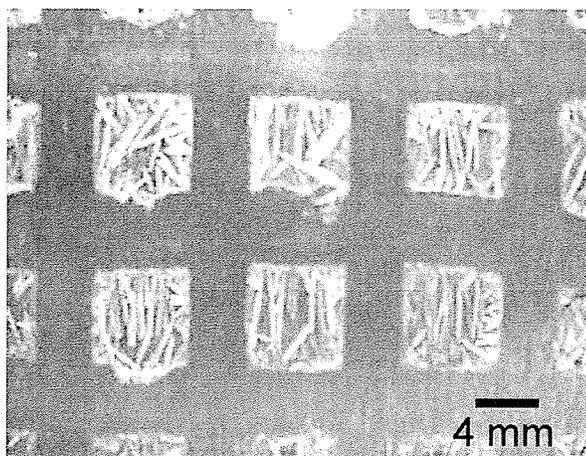


Figure 2.2: Hydrophobic sample barrier for MALDI-MS. Square spots prepared using the polybutadiene hydrophobic sample barrier. MALDI matrix (DHB) was pre-spotted using the tilt-and-tap technique. Peptides were deposited using the SepDep device.

2.3.7 Matrix Preparation, Tilt-and-tap pre-spotting

With rows and columns partitioned by hydrophobic polybutadiene, the MALDI target was tilted 60-70° from horizontal and 27 µL of matrix solution (DHB, 60 mg in 500 µL of 50:50 ACN/H₂O, 0.1% TFA) were added over the upper positions (1 and 2) of row A. The target was then lightly tapped onto the bench-top while maintaining the angle, causing the solution to run down its assigned row. Excess matrix solution was then allowed to run off the end of the target plate and wiped off with a damp lint-free paper tissue. The pre-spotting procedure was then repeated for each row (B through P). The estimated volume of matrix solution per MALDI sample spot was 0.8-1.0 µL. The square matrix spots were allowed to dry under ambient conditions prior to sample deposition.

2.3.8 Reverse phase microcolumns.

A column frit was made by carefully cutting the Fluoropore™ filter disk to approximately 0.25-0.50 mm² using a pair of tweezers and scissors. The small square frit was then transferred to the top-end of the Geloader™ pipette tip and carefully positioned in the lower portion of the column with the application of pressure using a 6 cm long, 290 µm o.d. steel capillary. Frit flow-through was tested by positioning the column on the device and applying a slight vacuum while adding a small volume of MeOH to the top-end of the column. Flow through the frit was adequate to clear 10-20 µL of MeOH within 20 seconds under full or near full vacuum. The column was then partially filled

with MeOH and bubbles were removed by gently tapping its side and/or by suction. Once all air bubbles were removed, C18 material in MeOH was added and allowed to partially settle before vacuum was re-applied to pack the stationary phase. Packing material was continually added until the final volume was between 3 – 5 μL . The column was then washed three times with 30 μL of 0.1% TFA using full vacuum. The final wash was allowed to reside approximately 2 mm above the LC bed for sample loading.

2.3.9 Sample loading and separation.

Eight 60 μL aliquots of solvents (10, 20, 30, 40, 50, 60, 70, and 80% ACN:H₂O, 0.1% TFA) were prepared. A peptide sample was loaded into the 2 mm of remaining wash solution. House vacuum was applied and the thumb valve was closed to obtain a chamber pressure of 0.3-0.5 atm, drawing the mobile phase through the column. Once the sample load resided ~0.5-1.0 mm from the top of the separating medium, vacuum was shut off and 5 μL of 0.1% TFA were added. Vacuum was then applied until the solution resided just above the chromatographic bed, followed by the addition of 45 μL of H₂O (0.1% TFA). The 45 μL solvent level was marked on the side of the column using a fine felt-tipped pen, followed by the addition of 15 μL of H₂O (0.1% TFA) to yield a total column volume of 60 μL . The target was then positioned within the chamber for the collection of the first fraction. As vacuum was established, eluent was drawn through the column and deposited on to the MALDI target. Repositioning of the device was achieved by reducing the

internal vacuum. As the solvent level reached the mark, 15 μ L of 10% ACN was added to the column until the full 60 μ L aliquot was consumed. As the deposition process continued, each solvent aliquot was added in 15 μ L increments, in order of increasing ACN concentration until elution was complete and the workflow outlined by Figure 2.3 was evaluated by MALDI-MS and MS/MS.

2.3.10 MALDI mass spectrometry.

Spectra were obtained on a Bruker Daltonics Biflex-IV instrument (Billerica, MA) with delayed extraction, operated in reflectron mode, using 16 V (lens 1) and 19 V (lens 2) acceleration voltages. The instrument was calibrated externally using angiotensin I, bombesin, substance P and ACTH 18-39 with target columns 1, 12 and 24 reserved for calibration. Spectra were acquired using the AutoExecute™ (Bruker Daltonics) interface with a maximum of 150 shots accumulated using 30-50 % relative laser power. Data were processed using M over Z™ software (www.genomicsolutions.com) with peak m/z values auto-labeled and saved to an Excel™ spread sheet for submission to Profound™ for peptide-mass fingerprinting (www.profound.com).

2.3.11 MALDI-MS/MS.

Once MALDI-TOF-MS spectra were analyzed, ions of interest were subjected to MS/MS sequencing. Spots on the 384 target (in DHB matrix)

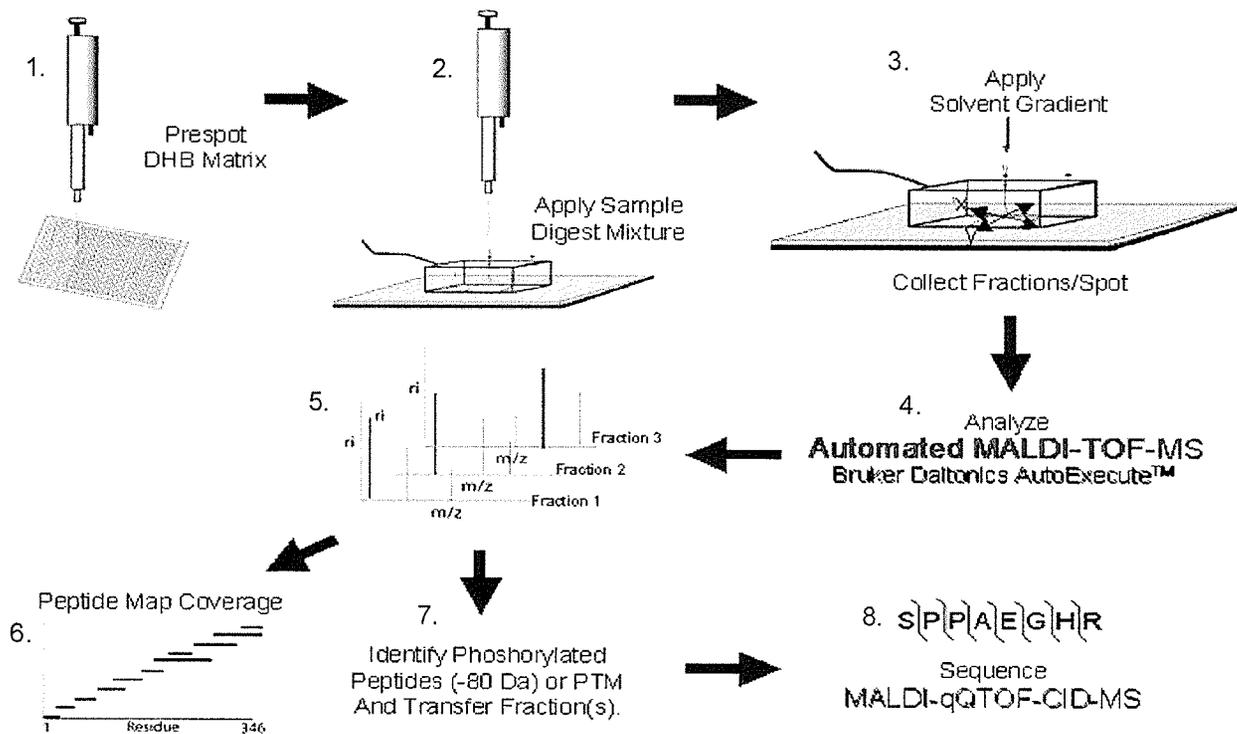


Figure 2.3: SepDep protocol. 1) DHB matrix is pre-spotted onto a MALDI target. 2) Protein digest is loaded onto the column. 3) Vacuum and solvent gradient are applied. 4,5) MALDI-TOF-MS is conducted by automated analysis and spectra are analyzed for peptide-mass mapping (6). 7) Masses corresponding to phosphorylated peptides are identified. 8) Fractions of interest are transferred using acetonitrile to a MALDI-QqTOF-MS/MS target for sequencing.

were dissolved in 1 μ L of 50% ACN in water. Solutions were mixed several times by aspiration into the pipette tip to ensure adequate transfer to a gold plated Intrinsic Bioprobes QqTOF target (Tempe, AZ). Peptides were then sequenced on an in-house built Manitoba/Sciex prototype MALDI-quadrupole/time-of-flight mass spectrometer (QqTOF) located at the University of Manitoba, Department of Physics and Astronomy (Loboda et al., 2000).

2.4 Results and Discussion

3.4.1 Design and performance of the SepDep device.

The overall performance of the SepDep device demonstrated that efficient peptide separations can be obtained using basic materials and a house vacuum line. The open column allows the application of a narrow-step gradient and the elution of peptides directly onto a pre-spotted MALDI target. This method improves MS peptide coverage by reducing or eliminating ion suppression effects.

As for design, some features of the device required consideration for adequate performance. For example, the thumb controlled release valve is positioned near a corner of the top panel to prevent disruption of the newly deposited sample spots by the returning atmosphere. This valve allows for the control of the column flow rate. Closed-cell weather stripping was used rather than a rigid rubber seal because of its compression characteristics that

aided sample deposition. Chamber pressures between 0.45 - 0.6 atm generated a drop of eluent (1.0-1.5 μL) in approximately 5-10 seconds, and higher vacuum (0.2 – 0.4 atm) resulted in full compression of the weather stripping. Device compression towards the target surface was effective in bringing the tip of the column (and eluent drop) to the MALDI surface without physical contact. This feature helped to reduce sample carry-over and clogging of the column by matrix that had been already deposited.

2.4.2 Polybutadiene hydrophobic sample barrier.

The polybutadiene hydrophobic barrier can be applied to any target surface in any desired geometry using the pen-type applicator. To date, no compatibility issues were encountered with the use of the polymer as described. The polymer is extremely hydrophobic in nature and very insoluble in water, methanol and acetonitrile solutions. As the target surface directly under the samples is not modified, target conductivity and charging properties are maintained. As a result, instrument settings such as acceleration voltage and ion optics require minimal re-optimization.

It should be noted that new Pap Pen markers are 'runny'. Sufficient care should thus be taken when applying the hydrophobic surface layer onto the target. Wiping excess material from the pen tip with a lint free tissue prior to application and using single strokes to partition each sample row and column leads to adequate results. It is also important to allow sufficient time

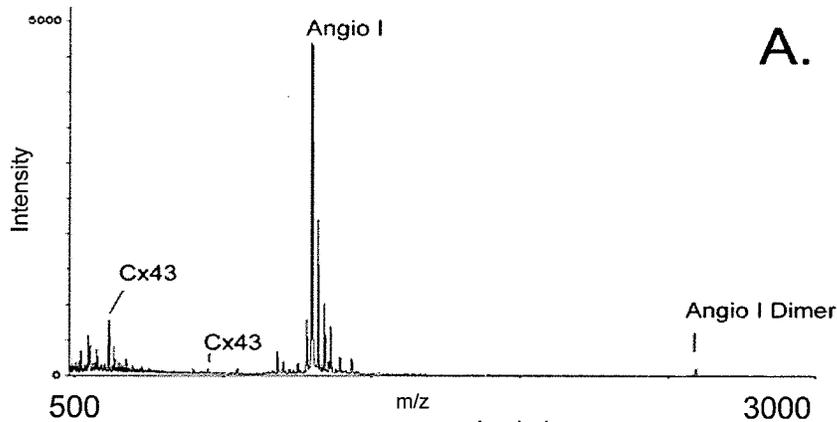
to cure newly treated targets, as Pap Pen solvents produce a significant amount of vapor. The barrier is re-usable after washes or rinses with 30% acetic acid/water or methanol. Removal of the hydrophobic barrier can be achieved using toluene and as described by product literature. With the hydrophobic barrier, more collected fractions produced useful MALDI spectra than with the traditional dried drop method. The barrier reduced edge effects and reduced the time that the automated acquisition routine took to find a 'sweet spot' (i.e. non-uniform areas of the crystalline sample containing higher proportions of peptides). The effectiveness of the barrier was especially noticeable with fractions eluting at higher ACN concentrations, i.e. with lower surface tension and spreading over a wider area on the target. The polybutadiene barrier was sufficiently hydrophobic to contain 60-75% ACN with volumes greater than 10 μL /spot.

The hydrophobic barrier also helps in the rapid application of matrix using the tilt-and-tap technique. All 384 target positions were easily pre-spotted with DHB solution and allowed to crystallize (Figure 2.2). The solubility characteristics of DHB in water and ACN resulted in the 1.0-1.5 μL of SepDep eluent nearly fully re-dissolving the matrix. Peptides were uniformly dispersed through the matrix when pre-deposited DHB was used, compared to less soluble matrices such as α -cyano-4-hydroxycinnamic acid.

The combination of Fluoropore™ column frit, column-packing material having a larger particle pore size and diameter than most HPLC media, plus the use of a house source vacuum, resulted in suitable flow rates for separation and deposition.

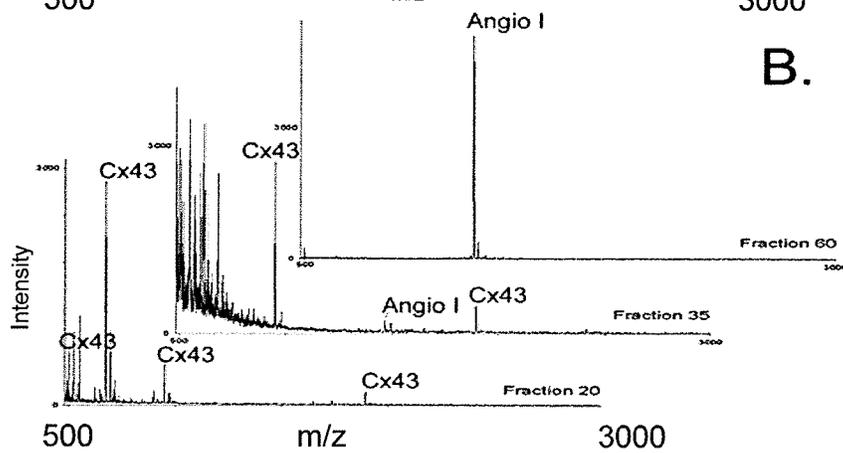
2.4.3 Reduction of ion suppression effects.

Cx43 peptides mixed with a greater than 19-fold molar concentration of angiotensin I (and sodium ions) underwent significant ion suppression. Only two of the theoretical four Cx43 peptides were detected as low abundance ions as seen in Figure 2.4A. Similar suppression effects were observed by Preisler *et al.* (2000) using an equimolar mixture of peptides with 50-fold excess of angiotensin II. After separation using the SepDep device, a significant reduction of suppression was achieved and all four theoretical Cx43 fragments were detected as abundant parent ions (Figure 2.4B). Tandem MS experiments were performed on ions of interest to certify the identity of the Cx43 derived peptides as shown in Figure 2.4C.



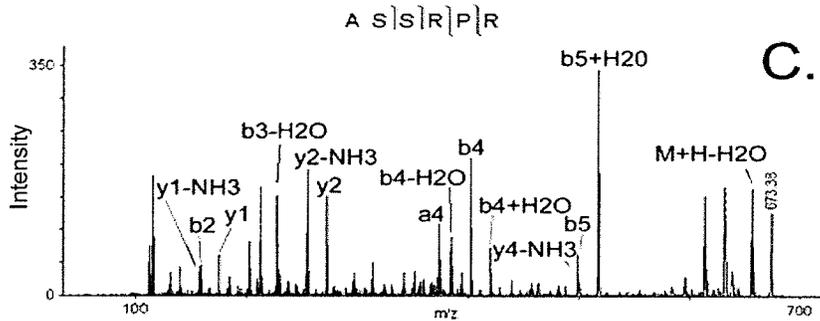
A. **Figure 2.4:** Reduction of MALDI ion suppression effects.

A) MALDI-MS showing suppression of Cx43 peptides by a 19 fold excess of angiotensin I peptide and Na⁺ ions. Only two out of the possible four Cx43 peptides can be observed.

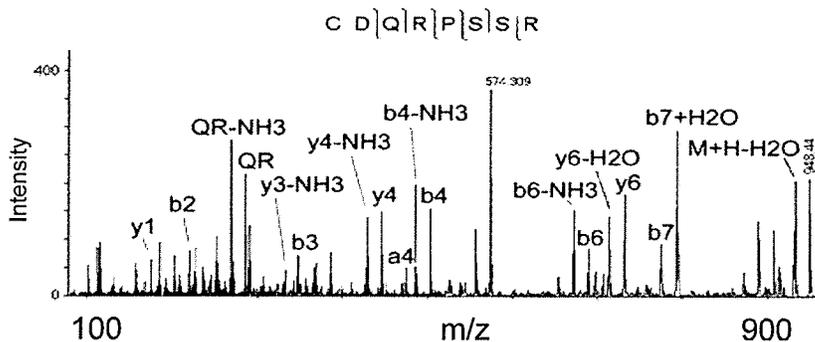


B. Sample partitioning using the SepDep device, and a narrow-step ACN/H₂O gradient resolving all Cx43 peptides.

B) Sample partitioning using the SepDep device, and a narrow-step ACN/H₂O gradient resolving all Cx43 peptides.



C. MS/MS of Cx43 peptides, *m/z* 948.44 and 673.38. Both spectra were obtained from fraction 25.



2.4.4 Separation of protein digests.

Following the successful separation of relatively simple peptide mixtures, the SepDep method was tested on tryptic peptides from larger proteins, *E. coli* citrate synthase (CS) and bovine β -casein. CS is a protein with 427 amino acids and a molecular weight of approximately 48 kDa. The sequence and structure of CS has been well characterized by X-ray crystallography (Nguyen et al., 2001; Maurus et al., 2003), and has been previously used to evaluate separation techniques used in preparation for mass spectrometry (McComb et al., 1999). The separation of CS tryptic fragments was achieved using a 0-80% ACN narrow-step gradient with the collection of 344 sample spots. Figure 2.5 shows the separation MALDI profile of 24 of the 344 spots that were analyzed by the auto-acquisition routine. Submission of m/z values to the Profound™ mass-fingerprint database led to the unambiguous verification of CS, with 52 matching peptides and minimum sequence coverage of 94% (Figure 2.6).

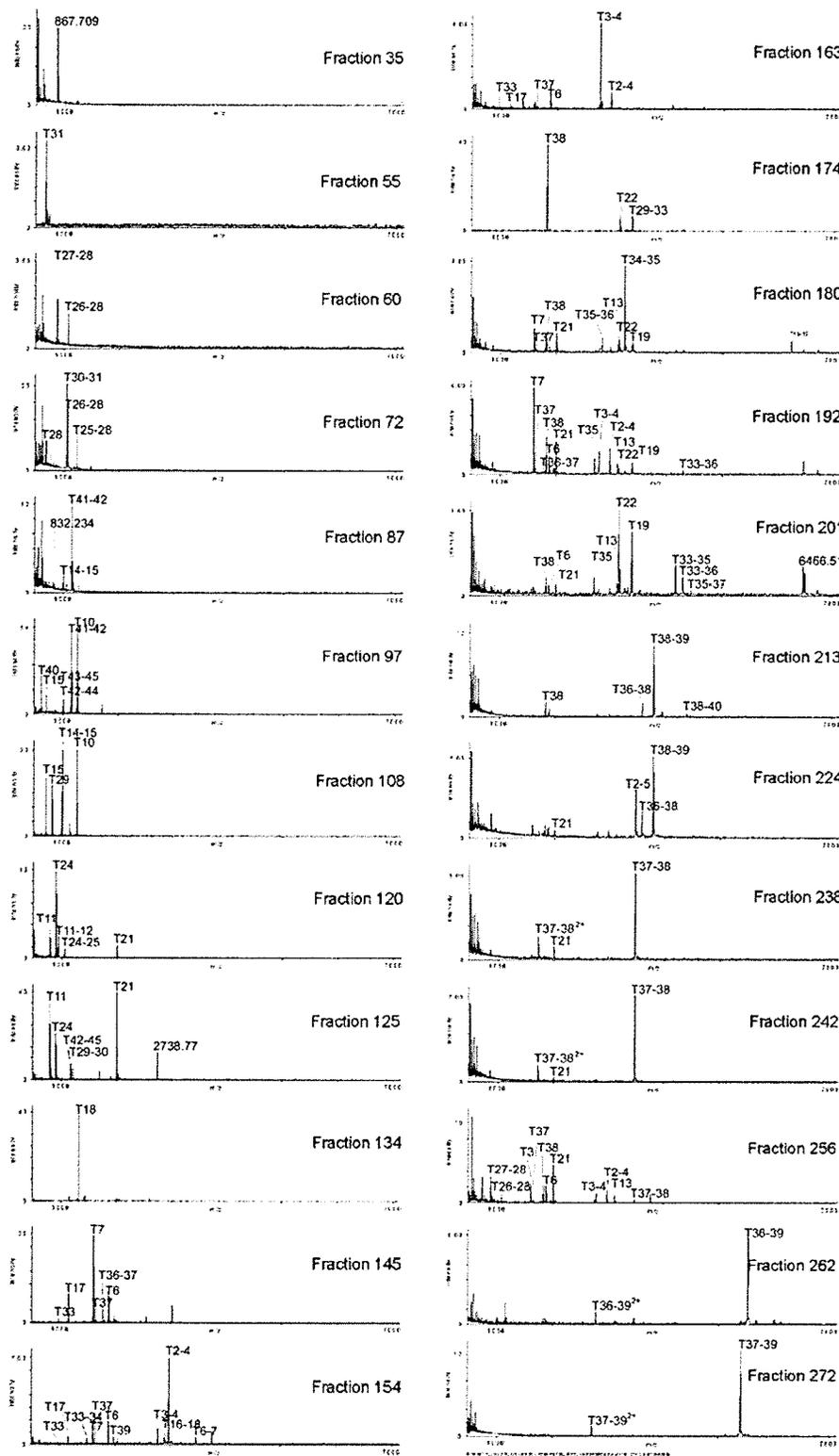


Figure 2.5: Representative MALDI-MS spectra from the off-line vacuum LC separation and deposition of citrate synthase tryptic peptides. Tryptic fragment (T#) and fraction number indicated.

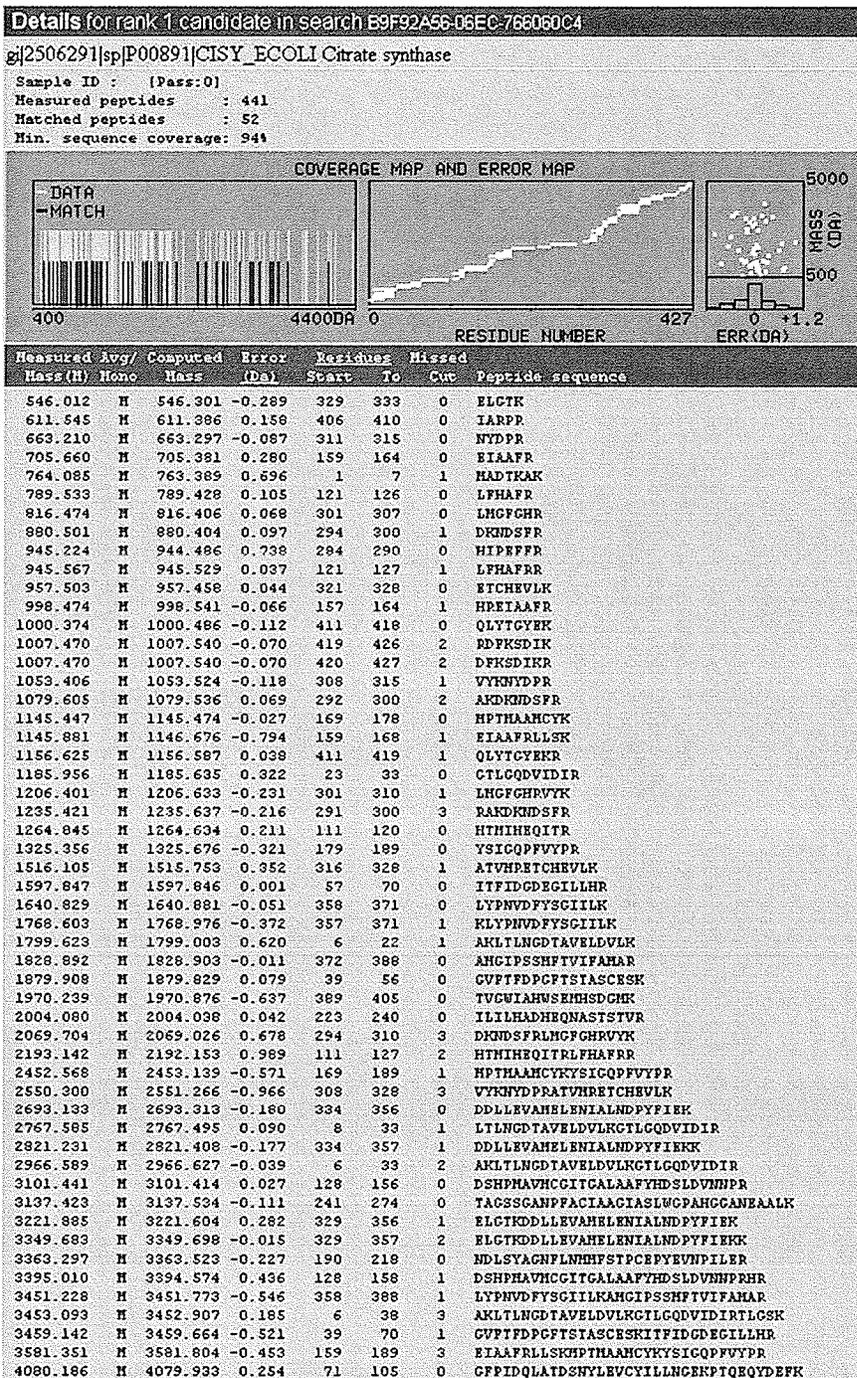


Figure 2.6: CS sequence coverage by SepDep MALDI-MS. Profound database search output demonstrating peptide mapping coverage of tryptic digest products within +/- 1 Da.

Separation of β -casein tryptic peptides was performed under similar conditions, and aimed at the analysis of phosphorylated peptides. β -casein is a penta-phosphorylated protein, with serines modified at positions 15, 17, 18, 19 and 35. Direct spotting of 2 pmole of β -casein digest using DHB yielded no detectable phosphorylated peptides. Additionally, negatively charged phosphopeptides are known to yield reduced ionization efficiencies (Figures 2.6A, 2.6C). Once separated, both predicted phosphorylated peptides, T1-2 (fraction 90, Figure 2.6B) and, T6 (fraction 31, Figure 2.7A) were detected as $[M+H]^+$ ions. Losses of HPO_3 moieties (80 u, by MALDI-TOF-MS, Figure 2.6B,) and H_2PO_3 (98 u, MALDI-QqTOF-MS) were also observed (Figure 2.7A, 2.7B). These losses were helpful to identify a non-specific phosphorylated peptide observed at m/z 2556 in fraction 60 (Figure 2.7D, 2.7E) for sequence analysis. MS/MS of m/z 2556 yielded FQpSEEQQTDELQDKIHPF (A.A.148-167) a non-tryptic peptide presumably due to the presence of chymotrypsin activity in the TPCK treated protease. Separation also resolved a second non-specific β -casein peptide corresponding to MHQPLPPTVMFPPQ (m/z 1981.9, Figure 2.9B, A.A. 159-175) observed in fraction 120, interestingly an isobar (i.e. distinct analytes having the same mass) of dephosphorylated T6 observed in fraction 31. Detection of both peptides corresponding to m/z 1981.9 and their respective sequencing and identification by MS/MS would not have been possible without separation of the peptide mixture.

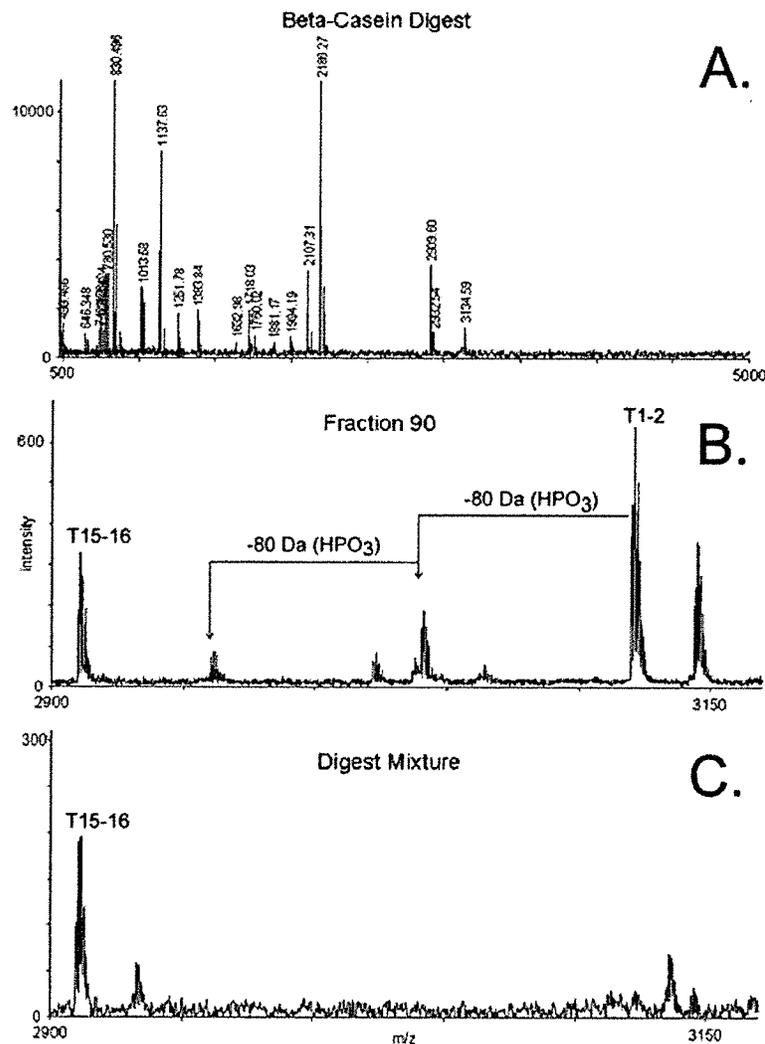
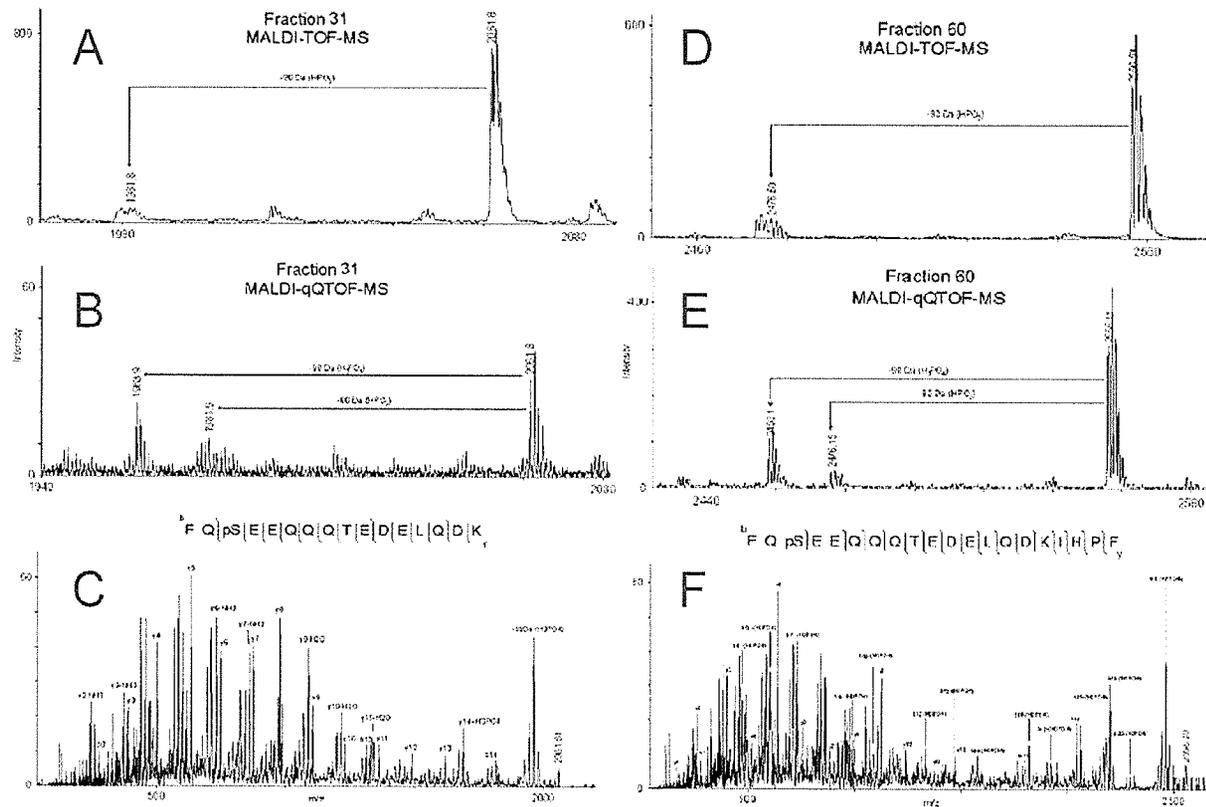


Figure 2.7: Identification of β -casein tetra-phosphorylated peptide. A) MALDI-TOF-MS of peptides derived from 2 pmole of phosphorylated β -casein. No m/z values corresponding to predicted phosphorylated peptides were observed by direct spotting. B) Identification of tetra-phosphorylated T1-2 (m/z 3123.98) and consecutive losses of HPO_3 (-80 Da) observed in fraction 90. C) Expanded region of the spectrum (shown in A) demonstrating no m/z values corresponding to phosphorylated or dephosphorylated T1-2.



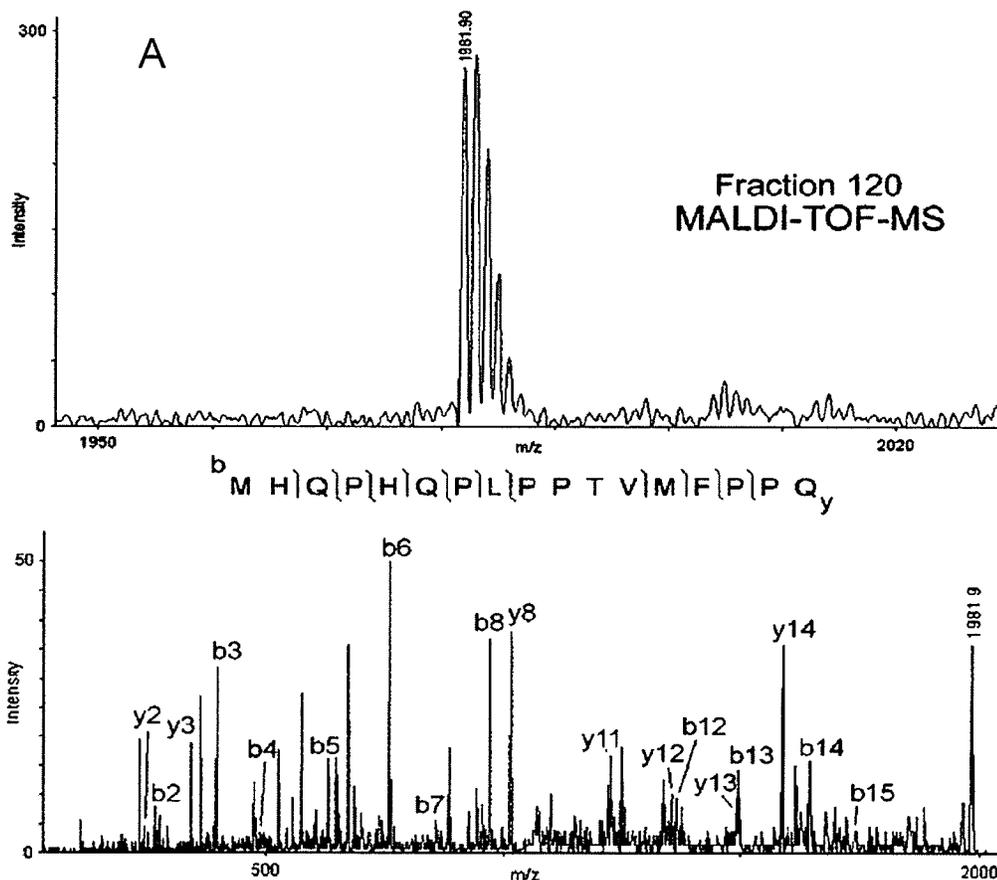


Figure 2.9: A) SepDep MS detection of β -casein peptide m/z 1981.9 (top) - having the same mass as dephosphorylated T6 observed in fraction 31. B) MS/MS and sequence analysis identified the peptide MHQPHQPLPPTVMFPPQ, a product from non-specific cleavage. Identification of both peptide isobars would have been difficult without chromatographic separation.

2.5 Conclusions

The SepDep device produces reasonable separations for MALDI analysis using simple, inexpensive materials and taking advantage of house vacuum. The technique significantly increases peptide coverage and

enhances MALDI spectral quality in general. The main drawbacks to this technique are the manual efforts necessary for separation and deposition. The incorporation of features such as an automated X-Y-Z stage and a solvent delivery/gradient system would decrease labour and would likely improve separation efficiency. In this report we demonstrated that a minimalist approach to peptide separation, the SepDep device, can be used to reduce ion suppression effects and increase the quality of peptide analysis by MALDI-MS.

2.6 Acknowledgements

We acknowledge Dr. Oleg Krokhin for acquiring preliminary QqTOF data, as well as Dr. Xinbo Li and Beverley Lee for helpful discussions. We also wish to thank Dr. H.W. Duckworth and members of his research group for helping to provide CS protein. The authors also acknowledge the Natural Sciences and Engineering Council of Canada (NSERC), the Canadian Foundation for Innovation (CFI), the Canadian Research Chair Program (CRC) and the National Health Institute of the USA (Grant No. GM-59240) to KGS for funding.

2.7 References

- Apffel, A., Chakel, J., Udiavar, S., Hancock, W. S., Souders, C., Pungor E. J. (1995) Application of capillary electrophoresis, high-performance liquid chromatography, on-line electrospray mass spectrometry and matrix-assisted laser desorption ionization--time of flight mass spectrometry to the characterization of single-chain plasminogen activator. *Chromatogr. A*, 717, 41-60.
- Asara, J.M, Alison, J. J. (1999) Enhanced detection of phosphopeptides in matrix-assisted laser desorption/ionization mass spectrometry using ammonium salts. *Am. Soc. Mass Spectrom.* 10, 35-44.
- Ayed, A., Krutchinsky, A.N., Ens, W., Standing, K.G., Duckworth, H.W. (1998) Quantitative evaluation of protein-protein and ligand-protein equilibria of a large allosteric enzyme by electrospray ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 12, 339-344.
- Bratt, C., Lindberg, C., Marko-Varga, G. (2001) Restricted access chromatographic sample preparation of low mass proteins expressed in human fibroblast cells for proteomics analysis. *J. Chromatogr. A*, 909, 279-288.
- Charlwood, J., Birrell, H., Tolson, D., Camilleri, P. (1998) Two-dimensional chromatography in the analysis of complex glycans from transferrin. *Anal. Chem.* 70, 2530-2535.
- Chen, V.C., Cheng, K., Lee, C., Ens, W., Standing, K.G., Nagy, J.I., Perreault, H., Am. Soc. For Mass Spectrometry, 51st ASMS Conference on Mass Spectrometry and Allied Topics, Montreal Canada, 8-16 June 2003: paper WPP279.
- Chen, V.C., Ens, W., Standing, K. G., Nagy, J. I., Perreault, H., HUPO 2nd Annual & IUBMB XIX Joint World Congress, Montreal, Canada, 8-11 October 2003: paper 31.26.
- Chong, B. E., Hamler, R. L., Lubman, D. M., Ethier, S. P., Rosenspire, A. J., Miller, F.R. (2001) Differential screening and mass mapping of proteins from premalignant and cancer cell lines using nonporous reversed-phase HPLC coupled with mass spectrometric analysis. *Anal. Chem.* 73, 1219-1227.
- Foret, F., Preisler, J. (2002) Liquid phase interfacing and miniaturization in matrix-assisted laser desorption/ionization mass spectrometry. *Proteomics*, 2, 360-372.
- Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., Roepstorff, P. (1999) Sample purification and preparation technique based on nano-scale reversed-

- phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrometry*, 34, 105-116.
- Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., Aebersold, R. (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. USA*, 97, 9390-9395.
- Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S., Millar, A., Taylor, P., Bennet, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B. D., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W. V., Figeys, D., Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*, 415, 180 – 183. Supplementary information.
- Hsieh, S., Dreisewerd, K., van der Schors R.C., Jiménez, C. R., Stahl-Zeng, J., Hillenkamp, F., Jorgenson, J. W., Geraerts, W. P. M., Li, K.W. (1998) Separation and identification of peptides in single neurons by microcolumn liquid chromatography-matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and postsource decay analysis. *Anal. Chem.* 70, 1847-1852.
- Krause, E., Wenschuh, H., Jungblut, P. R. (1999) The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins. *Anal. Chem.* 71, 4160-4165.
- Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C. et al. (2001) The Genome International Sequencing Consortium: Initial sequencing and analysis of the human genome. *Nature*. 409, 860-921.
- Loboda, A., Krutchinsky, A.N., Bromirski, M., Ens, W., Standing, K.G. (2000) A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance. *Rapid Commun. Mass Spectrom.* 14, 1047-1057.
- Mann, M., Wilm, M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. (1994) *Anal. Chem.* 66, 4390-4399.
- Mauri, P.L., Pietta, P.G., Maggioni, A., Cerquetti, M., Sebastianelli, A., Mastrantonio, P. (1999) Characterization of surface layer proteins from

- Clostridium difficile* by liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* 13, 695-703.
- Maurus, R., Nguyen, N. T., Stokell, D. J., Ayed, A., Hultin, P. G., Duckworth, H. W., Brayer, G. D. (2003) Insights into the evolution of allosteric properties. The NADH binding site of hexameric type II citrate synthases. *Biochemistry*, 42, 5555-5565.
- McComb, M.E., Donald L.J., Perreault H. (1999) Electrospray Ionization Mass Spectrometry and on-Line Capillary Zone Electrophoresis - Mass Spectrometry for the Characterization of Citrate Synthase, *Can. J. Chem.* 77, 1752-1760.
- Nguyen, N. T., Maurus, R., Stokell, D. J., Ayed, A., Duckworth, H. W., Brayer, G. D. (2001) Comparative analysis of folding and substrate binding sites between regulated hexameric type II citrate synthases and unregulated dimeric type I enzymes. *Biochemistry*, 40, 13177 –13187.
- Preisler, J., Hu, P., Rejtar, T., Karger, B.L. (2000) Capillary electrophoresis--matrix-assisted laser desorption/ionization time-of-flight mass spectrometry using a vacuum deposition interface. *Anal. Chem.* 72, 4785-4795.
- Preisler, J., Hu, P., Rejtar, T., Moskovets, E., and Karger, B. L. (2002) Capillary array electrophoresis-MALDI mass spectrometry using a vacuum deposition interface. *Anal. Chem.* 74, 17.
- Reid, G.E., Rasmussen, R. K., Dorow, D. S., Simpson, R. J. (1998) Capillary column chromatography improves sample preparation for mass spectrometric analysis: complete characterization of human alpha-enolase from two-dimensional gels following in situ proteolytic digestion. *Electrophoresis*, 19, 946-955.
- Rusconi, F., Schmitter, J. M., Rossier, J, Le Maire, M. (1998) Chromatographic separation and sample preparation in one step for MALDI mass spectrometric analysis of subpicomole amounts of heterogeneous protein samples. *Anal. Chem.* 70, 3046-3052.
- Udiavar, S., Apffel, A., Chakel, J., Swedberg, S., Hancock, W. S., Pungor, E. (1998) The Use of Multidimensional Liquid-Phase Separations and Mass Spectrometry for the Detailed Characterization of Posttranslational Modifications in Glycoproteins. *Anal. Chem.* 70, 3572-3578.
- Wall, D. B., Berger, S. J., Finch, J. W., Cohen, S. A., Richardson, K., Chapman, R., Drabble, D., Brown, J., Gostick D. (2002) Continuous sample deposition from reversed-phase liquid chromatography to tracks on a matrix-assisted

laser desorption/ionization precoated target for the analysis of protein digests. *Electrophoresis*. 23, 3193-3204.

Westman, A., Brinkmalm, G., Barofsky, D. F. (1997) MALDI induced saturation effects in chevron microchannel plate detectors. *Int. J. Mass Spectrom. Ion Proc.* 169, 79-87.

Zhang, H. Y., Stoeckli, M., Andren, P.E., Caprioli, R. M. (1999) Combining solid-phase preconcentration, capillary electrophoresis and off-line matrix-assisted laser desorption/ionization mass spectrometry: intracerebral metabolic processing of peptide E *in vivo*. *J. Mass Spectrom.*, 34, 377-383.

Part 3:

Development of proteomic approaches for the identification and
characterization of protein-protein interactions modulated by PDZ-domains

3.1 Author's Contributions

The original collaboration with the Nagy lab was established to identify connexin36 (Cx36) phosphorylation sites. Evidence provided at the time suggested that large quantities (>40 ng) of Cx36 protein could be obtained from GST-PDZ1 pull-downs using Cx36 over-expressing HeLa cells. Based on this information, it was hypothesized (X. Li and J.I. Nagy) and agreed upon (myself and H. Perreault) that the use of the PDZ1 would be suitable for the pull-down of Cx36 and phosphosite determination of the gap junction protein (V.Chen Ph.D Annual Review 2004). I then performed the pull-down experiments and related steps required to obtain (grow cells, induce expression), purify the GST-fusion protein and conjugate to agarose beads. Unfortunately the original aim of the study failed as the quality and quantity of Cx36 generated by the protocol were not suitable for proteomic analysis (V.Chen Ph.D. Annual Reviews 2004; 2005).

However as a number of proteins were clearly visible within SDS-PAGE gels, I generated and proposed the following analytical hypothesis: *the GST-PDZ1 could be used to enrich novel protein-protein interactions of ZO-1 for identification by MALDI-MS.* I then set out to identify other bands by solely conducting in gel digestions, sample preparation and MALDI analyses. Using the identities of the proteins and known binding characteristics of PDZs, I developed the PDZ screening process that was subsequently used to identify possible first order ZO-1 interactions, and the association of non-muscle α -

actinin with ZO-1. After evaluating the C-terminal sequence, I subsequently hypothesized that α -actinin directly interacts with ZO-1. At this time the findings were communicated to H. Perreault, X. Li and J.I. Nagy, and it was agreed that the findings were of merit for further investigation.

Due to the sequence similarity between α -actinin-1 and α -actinin-4, I established the identity of the actinin by MS/MS. Identity of the α -actinin was necessary as protein specific commercial antibodies were required, leading to the acquisition of the polyclonal anti- α -actinin-4 Alexis Cat. No. 210-356-C050. Upon these positive results, it was recognized that a larger quantity of anti-actinin-4 would be needed to perform experiments. Due to this requirement, a new antibody (42-1400) was developed by the Nagy lab in collaboration with Zymed/Invitrogen. Immunoblots confirming the specificity of the new Zymed/Invitrogen antibody, in particular those found within Figures 3.2 and 3.3, were performed by X. Li.

Preliminary co-localization experiments using HeLa/Cx36 transfected HeLa and mouse liver tissues were performed by me under the direction of X. Li and some of the images found within Figures 3.6 and 3.7 were acquired by Mr. B. McLean under the direction of J.I. Nagy, with corresponding figure captions also provided by J.I. Nagy. Immunoblots found in Figure 3.3 and 3.4 were performed by myself in collaboration with X. Li. Experiments establishing that only non-muscle α -actinin-4 (not α -actinin-1) was capable of

associating with ZO-1, as presented by Figure 3.5, were designed/performed by me and on record within lab journals.

In terms of the manuscript, I was responsible for literature review and data interpretation leading to the first series of manuscript drafts. I also formulated a mechanism of transcriptional regulation based on the interaction of α -actinin-4 with ZO-1 (section 3.5.3 Functional Implications; V. Chen PhD. Annual Reviews 2005 and 2006). Final drafts were revised in consultation with J.I. Nagy. All written materials and communications related to the Journal of Proteome Research publication, including letter to the editor were drafted by me, and edited by J.I. Nagy prior to submission. The final draft of this manuscript was reviewed, corrected and approved by all listed authors.

-V. Chen

Interaction of zonula occludens-1 (ZO-1) with α -actinin-4: application of functional proteomics for identification of PDZ domain-associated proteins

Vincent C. Chen¹, Xinbo Li², H el ene Perreault¹, and James I. Nagy²

1. Department of Chemistry, University of Manitoba
Winnipeg, Manitoba, Canada

2. Department of Physiology, University of Manitoba
Winnipeg, Manitoba, Canada

Published: Journal of Proteome Research, 5 (9), 2123 -2134, 2006.

*Reproduced with permission from *Journal of Proteome Research* 5 (9), 2123 -2134, 2006.

Copyright   2006 American Chemical Society

3.1 Abstract

The use of recombinant “bait” proteins to capture protein binding partners followed by identification of protein interaction networks by mass spectrometry (MS) has gained popularity and widespread acceptance. We have developed an approach using recombinant PDZ protein interaction modules of the membrane-associated guanylate kinase (MAGUK) protein zonula occludens-1 (ZO-1) to pull-down and screen for proteins that interact with these modules via their PDZ domain binding motifs. Identification of proteins by MS of pull-down material was achieved using a vacuum-based chromatography sample preparation device designed for matrix-assisted laser desorption/ionization (MALDI) MS. MS analysis of tryptic fragments in pull-down material revealed a number of potential ZO-1 interacting candidates, including the presence of peptides corresponding to the cortical membrane scaffolding protein α -actinin-4. Interaction of α -actinin-4 with ZO-1 was confirmed by co-immunoprecipitation of these two proteins from cultured cells as well as from brain, liver and heart, and by immunoblot detection of α -actinin-4 after pull-down with the first PDZ domain of ZO-1. In contrast, the highly homologous α -actinin family member, α -actinin-1, displayed no association with ZO-1. Immunofluorescence showed co-localization of α -actinin-4 with ZO-1 in cultured HeLa and C6 glioma cells, as well as in a variety of tissues *in vivo*, including brain, heart, liver and lung. This study demonstrates the utility of MS-based functional proteomics for identifying cellular components of the ZO-1 scaffolding network. Our finding of the

interaction of ZO-1 with α -actinin-4 provides a mechanism for linking the known protein recruitment signaling activities of ZO-1 with α -actinin-4-associated plasma membrane proteins that have regulatory activities at cell-cell and cell-extracellular matrix contacts.

3.2 Introduction

Zonula occludens-1 (ZO-1) belongs to a class of plasma membrane-associated guanylate kinase scaffolding proteins (MAGUK) that contain guanylate kinase, Src homology-3 (SH3) and multiple PDZ domains (post-synaptic density-95, *drosophila melanogaster* discs large, ZO-1). ZO-1 is a component of a variety of intercellular junctions, where it interacts with a wide range of transmembrane and cortical membrane proteins, including claudins and occludins at tight junctions (Itoh et al., 1999; Mitic et al., 1999), α , β and γ catenins at adherens junctions (Itoh et al., 1993; Yokoyama et al., 2001; Rajasekaran et al., 1996), various connexins (Cx31.9, Cx36, Cx43, Cx45, Cx46, Cx47, Cx50) at gap junctions (Giepmans et al., 1998; Li et al., 2004^a; Kausalya et al., 2001; Laing et al., 2001; Nielsen et al., 2002) and nerve terminal proteins at synaptic junctions (Inagaki et al., 2003). In addition, ZO-1 contains multiple nuclear import signals and was recently reported to interact with the Y-box transcription factor ZONAB (Gottardi et al., 1996; Islas et al., 2002; Balda et al., 2000; Balda et al., 2003), suggesting that it shuttles between the nucleus and the plasma membrane. Multiple PDZ domain-containing proteins serve as signaling and scaffolding hubs that localize

distinct but functionally related proteins at the inner surface of the plasma membrane. Such a role of ZO-1 is indicated by its widespread presence at a number of different cellular structures. Thus, identification of additional ZO-1 interacting partners is relevant to a broad range of physiological processes.

PDZ domains are peptide binding pockets that are formed by six β -strands, two α -helices and a conserved 'R/K-X-X-X-G-L-G-F carboxylate binding loop', and that generally interact with the last four to six c-terminus amino acids of transmembrane and membrane scaffolding proteins. Nomenclature of PDZ domain binding ligands designate the last amino acid as P_0 , with subsequent residues assigned as P_{-1} , P_{-2} , P_{-3} etc. The P_0 and P_{-2} residues are critical for PDZ ligand recognition, although some variability has been observed (Songyang et al., 1997). Three main classes of PDZ domain binding ligands have been identified based on their preference for interaction with PDZ domains; class I PDZ binding ligands contain the sequence S/T-X- Φ (where Φ is a hydrophobic amino acid), class II ligands contain Φ -X- Φ , and class III ligands contain D-X-V.

In recent years, peptide-mass fingerprinting involving mass measurement of peptides generated after proteolytic digestion and tandem-MS peptide fragmentation analysis have become powerful tools to identify and study protein interaction networks. This approach has been largely based on the use of recombinant 'bait' proteins to capture binding partners, followed

by electrophoretic (SDS-PAGE) or liquid chromatographic separation (LC) for identification of captured proteins by mass spectrometry (MS). Compared to technologies that are used to study protein interactions such as yeast-two-hybrid screens, MS-based studies are effective in deciphering components within large protein interaction complexes and networks. Here, we present a method for identifying PDZ domain-mediated protein interaction networks using a discrete PDZ domain to pull-down interacting proteins, which were then identified by MALDI-MS. Full sequences of identified proteins were evaluated for the presence of class I, II and III PDZ c-terminus binding motifs, which were then used to generate a list of protein interaction candidates of ZO-1. Among these candidates, this approach revealed a novel protein interaction between ZO-1 and a member of α -actinin family of proteins, namely α -actinin-4. Biological relevance of this interaction was demonstrated by pull-down assay, co-immunoprecipitation and immunofluorescence co-localization of ZO-1 with α -actinin-4 in cultured cells and mouse tissues. The method outlined demonstrates MS-based functional proteomics as a hypothesis-generating tool for the discovery and characterization of protein complexes, and in particular those protein-protein interactions mediated by PDZ modules, and provides evidence that α -actinin-4 is a functionally distinct member of the non-muscle α -actinin family.

3.3 Experimental Procedures

3.3.1 Antibodies and Animals.

Antibodies against α -actinin-4, α -actinin-1, ZO-1 and anti-GST were used in this study. Rabbit anti- α -actinin-4 antibodies were obtained from ALEXIS USA (Cat. No. 210-356-C050; San Diego, CA, USA) and from Invitrogen/Zymed Laboratories (Cat No. 42-1400; South San Francisco, CA, USA). Monoclonal antibody specific to non-muscle α -actinin-1 was obtained from Sigma (clone BM 75.2, Cat. No. A5044, St. Louis, MO, USA). Monoclonal anti-ZO-1 (Ab33-9100) and polyclonal anti-ZO-1 (Ab61-7300) were obtained from Invitrogen/Zymed Laboratories, and rabbit anti-GST (Ab06-332) used for detecting GST fusion proteins was obtained from Upstate Cell Signaling Solutions (Lake Placid, NY, USA). All secondary antibodies used for immunofluorescence were obtained from Jackson ImmunoResearch Laboratories (West Grove, PA, USA). A total of fifteen adult male CD1 mice were used according to protocols approved by the Animal Care Committee at our institution, with minimization of stress to animals. All chemicals and reagents described were obtained from Sigma unless otherwise stated.

3.3.2 Cell culture and tissue preparation.

HeLa cells (American Type Culture Collection, Rockville, MD, USA), HeLa cells stably transfected with connexin36 (Cx36) and C6 glioma cells were grown in Dulbecco's modified Eagle's medium supplemented with 10%

fetal bovine serum and 1% penicillin-streptomycin (Appendix: Table 8.6). We previously used HeLa cells transfected with Cx36 to establish Cx36 interaction with ZO-1, and these cells served as a source of tissue for some of the present studies (Li et al. 2004). Cultured cells were rinsed briefly with PBS (50 mM sodium phosphate buffer, pH 7.4, 0.9% saline) and lysed using immunoprecipitation (IP) buffer containing 20 mM Tris-HCl, pH 8.0, 140 mM NaCl, 1% Triton X-100, 10% glycerol, 1 mM EGTA, 1.5 mM MgCl₂, 1 mM dithiothreitol, 1 mM phenylmethylsulphonyl fluoride, 5 µg/mL and protease inhibitor cocktail consisting of leupeptin, pepstatin A and aprotinin. Tissues from mice were collected, rapidly frozen and stored at -80 °C. Tissues were disrupted in homogenization buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 150 mM NaCl, 1% Triton-X 100, 1 mM sodium orthovanadate, 1 mM phenylmethylsulphonyl fluoride and 2 µg/mL pepstatin A, leupeptin and aprotinin) and taken for pull-down, IP and western blotting.

3.3.3 GST-PDZ domain fusion proteins.

pGEX-3X plasmids each containing one of the three PDZ domains of ZO-1 linked to GST were kindly provided by Dr. Ben Giepmans (University of California, San Diego, CA, USA). GST-PDZ domain fusion protein expression in *Escherichia coli* DH5α and attachment of the domains to glutathione-Sepharose 4B beads for pull-down assays were conducted using standard procedures as previously described (Nielsen et al., 2002, Nielsen et al., 2003;

Li et al., 2005). All fusion proteins were subsequently characterized by MALDI-MS as described below.

3.3.4 Pull-down of proteins for MS analysis.

Wild-type HeLa cells, HeLa cells stably transfected with Cx36 and C6 glioma cells were used to screen for proteins that interact with the PDZ1 domain of ZO-1. These cells were grown to 90% confluence, harvested in IP buffer and collected (Appendix: Table 8.5 and 8.6) in tubes each containing 150 mg of cells in 750 μ L and sonicated. Cellular debris was removed by centrifugation at 20,000 g for 10 minutes at 4 $^{\circ}$ C. Supernatants were pre-cleared for 1 hour at 4 $^{\circ}$ C with 75 μ L of glutathione-linked Sepharose beads to remove proteins interacting non-specifically with the beads, and supernatants were collected after centrifugation at 20,000 g at 4 $^{\circ}$ C. *In vitro* binding was performed by incubation of 60 μ L of Sepharose-conjugated GST-PDZ1 with cellular supernatants for 16 hours at 4 $^{\circ}$ C. Beads were then collected, washed twice for 20 minutes at 4 $^{\circ}$ C with 1 mL of washing buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.5% NP-40) and incubated for 1 hr with 200 μ L of 500 μ g/mL of peptide (P-N-F-G-R-T-Q-S-S-D-S-A-Y-V-COOH, Tufts Protein Chemistry Facility, Tufts University) corresponding to the C-terminus sequence of Cx36. This peptide as well as full-length Cx36 is known to interact directly with the PDZ1 domain of ZO-1, and peptide at a concentration of 100 μ M was previously found to produce a large reduction in ZO-1 interaction with Cx36 (Li et al., 2004^a; Li et al., 2004^b). Proteins eluted

from the beads with peptide were concentrated to approximately 40 μL using a 10,000 MWCO Microcon™ (Millipore, Billerica, MA) centrifuge filtration device. Concentrated proteins were then treated with 40 μL of SDS-PAGE loading buffer containing 10% of β -mercaptoethanol and briefly heated at 60 °C for 2 minutes prior to loading. The loading solution was then evenly distributed over 3 sample wells, and separated electrophoretically in a 12.5% SDS-PAGE gel for 60 minutes at 200 V using a Bio-Rad Mini-Protean II™ and Power Pac™ (Hercules, CA) as described in Appendix: Table 8.0.

3.3.5 Mass Spectrometry.

After SDS-PAGE separation, proteins were visualized with Gelcode™ colloidal Coomassie blue stain (Pierce, Rockford, IL, USA). Segments of gel containing bands corresponding to the same molecular weight across each of the three sample lanes were removed and pooled. The gel segments were destained, digested with sequence-grade modified trypsin (Promega, Madison, WI), and extracted using standard in-gel digestion procedures (Appendix: Table 8.3). Peptide extracts were dried under vacuum using a Savant Speedvac™ (Farmingdale, NY, USA), resuspended in 5 μL of 0.1% trifluoroacetic acid in water, and divided into 1 μL and 4 μL portions for solid phase extraction using C18 Ziptips™ (Millipore) (Appendix: Table 8.4) and reverse phase (C18) μ -VLC separation as described below and within Part 2 of this thesis. Extracted peptides were mixed with an equal volume of 2,5-

dihydroxybenzoic acid matrix (120 mg/mL in 1:1 acetonitrile/water with 0.5% of trifluoroacetic acid) for MS analysis.

Peptide-mass fingerprints were acquired on a Bruker BiflexTM IV (Billerica, MA) Reflectron time-of-flight (TOF) instrument operated in the delayed extraction mode. Peptide monoisotopic peaks with a signal-to-noise ratio of 1.6 or greater were assigned using *M over Z*TM software (Genomic Solutions, Ann Arbor, MI), and resulting values were submitted to the *Mascot*TM protein identification server (Matrix Science, www.matrixscience.com) using the NCBI non-redundant (NCBI nr) database.

All MS/MS sequencing was performed on digest mixtures chromatographically separated using a custom built, off-line separation and on-target deposition device developed in-lab (Part 2, Chen et al., 2004). The method involves use of micro-columns packed with C18 reverse-phase stationary material (J.T. Baker, 15 μ m, 300 Å) for separation of proteolytic mixtures with a 0-80% acetonitrile/water (0.1% trifluoroacetic acid) solvent gradient. Localization of sample deposits on stainless steel targets was aided by modifying the MALDI surface with a hydrophobic polybutadiene sample barrier (PapPen^{HT}, Diado, Japan) as previously described (Part 2, section 2.3.6; Chen et al., 2004). MALDI spots containing peptides of interest identified in MS mode were dissolved in 2 μ L of 0.1% trifluoroacetic acid in water and transferred to a custom manufactured gold-plated target (Intrinsic

Bioprobes Inc., Tempe, AZ) for sequencing on a Manitoba/Sciex prototype hybrid MALDI quadrupole/time-of-flight (QqTOF) mass spectrometer (Loboda et al., 2000) MS/MS spectra were annotated using *M over Z*TM (Genomic Solutions, www.genomicsolutions.com) as previously summarized in Figure 2.3. Mass-to-charge ratios (*m/z*) of fragments were submitted to *SonarMS/MS*TM (Genomic Solutions, www.genomicsolutions.com) using the NCBI nr and manually evaluated with resources available from Protein Prospector (<http://prospector.ucsf.edu/>) using the human sequence for α -actinin-4 (gi:12025678).

3.3.6 Pull-down for confirmation of PDZ/ α -actinin-4 interaction.

Recombinantly expressed GST fusion proteins containing either the PDZ1, PDZ2 or PDZ3 domains of ZO-1 were prepared and conjugated as previously described (Li et al., 2003^a) and characterized by MALDI-MS (Figure XX). Sepharose-conjugated fusion proteins were incubated with cell lysates in IP buffer for 16 hours at 4 °C. Beads were then washed 5 times in PBS buffer containing 1% Triton X-100. GST-PDZ associated proteins were then eluted with SDS-PAGE sample buffer and taken for immunoblotting with anti- α -actinin-4. Immunoblot membranes were stripped (with stripping buffer Appendix: Table 8.10) and reprobed with anti-GST antibody diluted at 1:10,000 to confirm equal loading of the fusion proteins.

3.3.7 Immunoprecipitation (IP).

HeLa cells, C6 cells and samples of brain, liver and heart were homogenized by brief sonication in IP buffer. Cellular debris was removed by centrifugation at 20,000 g for 10 minutes at 4°C, and 2 mg of supernatant protein was washed for 1 hour at 4 °C with 20 µL of protein A-coated agarose beads (Santa Cruz BioTech, Santa Cruz, CA, USA). Samples were then centrifuged at 20,000 g for 10 minutes at 4 °C, and supernatants were incubated with 2 µg of polyclonal anti-ZO-1 antibody with shaking for 2 hours at 4 °C, followed by 1 hour of incubation with 20 µL of protein-A coated agarose beads. The mixture was then centrifuged at 20,000 g for 10 minutes at 4 °C, and pelleted beads were washed five times under rotation with 1 mL of washing buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl and 0.5% NP-40). The beads were then collected and boiled for 3 minutes in SDS-PAGE loading buffer containing 10% β-mercaptoethanol and taken for immunoblot detection of α-actinin-4 or α-actinin-1 (Appendix: Tables 6.0 and 6.1). Control samples were taken through the IP procedure with exclusion of primary antibody. Some membranes were washed in stripping buffer (2% SDS, 100 mM β-mercaptoethanol, 65.5 mM tris-HCl, pH 6.7) for 30 minutes at 60 °C and reprobbed with monoclonal α-actinin-1 or polyclonal α-actinin-4 antibody, respectively, for western blot processing as above.

3.3.8 Western blotting.

Cell lysates, tissue homogenates and IP material were sonicated and briefly centrifuged, followed by protein determination using a kit (Bio-Rad Laboratories, Hercules, CA, USA). Sample proteins (20 µg) were separated electrophoretically in 12.5% polyacrylamide gels and transblotted to polyvinylidene difluoride membranes (Bio-Rad) in standard Tris-glycine transfer buffer (pH 8.3) containing 0.5% SDS. Membranes were blocked for 2 hours at room temperature in TSTw (20 mM Tris-HCl, pH 7.4, 150 mM NaCl with 0.2% Tween-20) containing 5% non-fat milk powder, followed by a brief wash with TSTw. Membranes were incubated for 16 hours at 4 °C with anti-actinin-4 or α -actinin-1 antibody (0.5 µg/mL) in TSTw containing 1% non-fat milk powder, then washed four times in TSTw for 40 minutes, and incubated with horseradish peroxidase-conjugated donkey anti-rabbit IgG diluted 1:5,000 (Sigma-Aldrich Canada, Oakville, ON, Canada) in TSTw containing 1% non-fat milk powder. Membranes were finally washed with TSTw four times for 40 minutes and resolved by chemiluminescence (ECL, Amersham PB, Baie d'Urfe, Quebec, Canada).

3.3.9 Light microscope immunofluorescence.

HeLa cells and C6 glioma cells cultured on poly-L-lysine-treated glass cover slips were rinsed with PBS and fixed in ice-cold 1 or 2% formaldehyde for 5 minutes (Appendix: Table 8.6). Fixed cells were incubated for 16 hours at 4 °C with polyclonal anti- α -actinin-4 antibody at a concentration of 1 µg/mL

and monoclonal anti-ZO-1 antibody at 0.5 µg/mL in TBST buffer (50 mM Tris-HCl, pH 7.6, 1.5% NaCl, 0.3% Triton X-100) containing 4% normal donkey serum. Cells were then washed four times with TBST for 1 hour, incubated for 1 hour simultaneously with fluorescein isothiocyanate (FITC)-conjugated donkey anti-rabbit IgG diluted at 1:200 and Cy3-conjugated goat anti-mouse IgG diluted at 1:200, or alternatively with FITC-conjugated horse anti-mouse IgG diluted 1:200 and Cy3-conjugated donkey anti-rabbit IgG diluted 1:200 (Table 8.7). Cover slips were then washed for 1 hour in TBST and mounted on glass slides with antifade medium (Table 8.9).

Mice were deeply anaesthetized with equithesin and transcardially perfused sequentially with cold (4 °C) solutions of prefixative (PBS containing 0.1% sodium nitrite and 1 unit/mL of heparin), then fixative (0.16 M sodium phosphate buffer, pH 7.6, 0.2% picric acid and 1 or 2% paraformaldehyde) and then a sucrose flush (10% sucrose, 25 mM sodium phosphate buffer, pH 7.4). Tissues were removed and stored in cryoprotectant (25 mM sodium phosphate buffer, pH 7.4, 10% sucrose, 0.04% sodium azide) for 24 hours. Cryostat sections (10 µm thick) of tissues collected on gelatinized glass slides were processed for double immunofluorescence labeling with all antibodies diluted in TBST containing 10% normal donkey serum. Sections were incubated for 16 hours at 4 °C simultaneously with monoclonal anti-ZO-1 (3 µg/mL) and polyclonal anti- α -actinin-4 (1 µg/mL), then washed for 1 hour in TBST, and incubated for 1.5 hours at room temperature with FITC-conjugated

horse anti-mouse IgG diluted 1:200 and Cy3-conjugated donkey anti-rabbit IgG diluted 1:400, or Cy3- conjugated goat anti-mouse IgG diluted 1:400 and FITC-conjugated donkey anti-rabbit IgG diluted 1:200. Sections were then washed for 1 hour in TBST and coverslipped with anti-fade medium (as outlined by Table 8.8). Immunofluorescence was examined on a Zeiss Axioscop 2 fluorescence microscope. Images were obtained using Carl Zeiss Axiovision 3.0 image software (Carl Zeiss Canada, Toronto, ON) and assembled using Northern Eclipse 5.0, Photoshop 3.0 and Corel Draw 8 software.

3.4 Results

3.4.0 Characterization of recombinant fusion proteins by MALDI-MS

To confirm recombinant expression, proteins were separated by SDS-PAGE and removed by in-gel digestion. Resulting MS spectra (Figure 3.0A) were analyzed and proteins were identified by PMF, confirming the presence of the GST-fusion proteins (glutathione-S-transferase (Figure 3.0B) and ZO-1 (Figure 3.0C). High precision and accuracy of the instrument (MALDI-QqTOF described in subsequent sections) led to the identification of ZO-1 with only 2% sequence coverage.

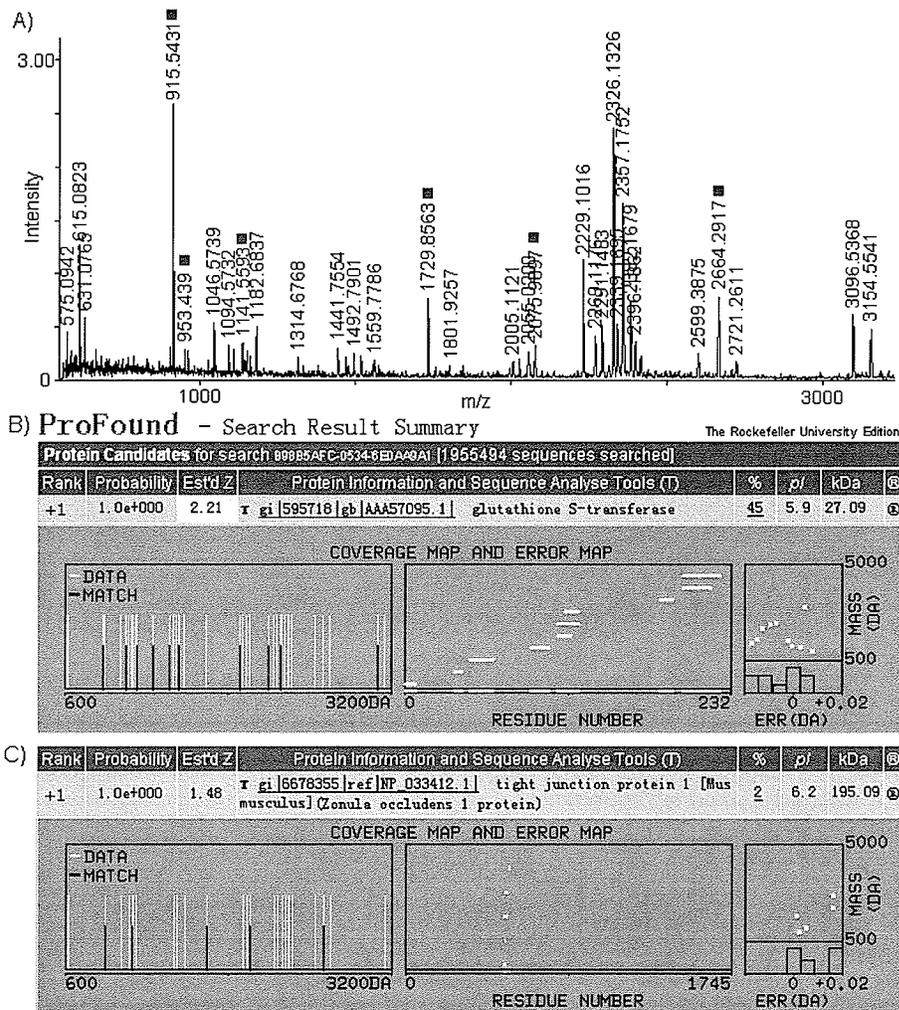


Figure 3.0: Analysis of ZO-1 GST-fusion proteins by MALDI-MS. (A) The GST-SH3 conjugated to glutathione-Sepharose beads, washed and eluates were purified by SDS-PAGE for in-gel tryptic digestion and MALDI-MS. Generated mass lists were then submitted to a PMF search engine (Profound™, <http://prowl.rockefeller.edu>). Peaks corresponding to ZO-1 have been highlighted with red squares. (B, C) Database search (Profound) confirming the identities of the GST and tight junction protein-1 (ZO-1) within the recombinant sample. ZO-1 sequence coverage corresponds to the region of the protein which was recombinantly expressed within the GST-fusion protein.

3.4.1 MS identification of α -actinin-4 pull-down with the PDZ1 domain of ZO-1.

HeLa cell lysates were initially used to search for new ZO-1 binding proteins by a proteomic-based approach involving pull-down with a fusion protein consisting of GST linked to the PDZ1 domain of ZO-1. The competitive elution of PDZ1-associated proteins using a peptide corresponding to the c-terminal tail of Cx36 resulted in the recovery of several protein bands that were observed after Coomassie staining (Figure 3.1A). Proteins that were identified by peptide-mass fingerprinting and that contained putative PDZ domain class I, II or III interaction motif sequences were compiled and summarized (Table 3.1), with α -actinin selected for further study. Contrary to expectation, we were unable to identify Cx36 by MS in pull-down samples from cells expressing Cx36. This was likely due to the presence of an abundant protein migrating at 36 kDa, which was identified as the enzyme carbonyl reductase. This enzyme is known to bind covalently to GST via a cysteine residue in its active site, and its large quantities in gel digests effectively masked MS detection of Cx36, which was confirmed to be present as previously described (Li et al 2004^a; Li et al 2004^b).

Table 3.1: Identification of ZO-1-associated scaffolding network by MALDI-MS.

Identified PDZ1 pull-down proteins (HeLa), GI accession	Mascot Score	C-terminal region	Putative C-terminal PDZ class	Known associations	Primary Cellular localization
α-actinin, non-muscle 3157976	110			Actin, Integrins ($\beta 1$, $\beta 2$, $\beta 3$), ICAMs, α -catenin, BP180, MAGI-1, ICAM-1, ICAM-2, Vinculin, ADIP, L-selectin, CRP (Otey et al., 2004)	Plasma membrane-associated
actinin, alpha 4 13477151	110	E-S-D-L-COOH	Class I		
ACTB protein 15277503	98			Many proteins including: Plectin, α -actinin, Vimentin, α -Spectrin/Fodrin (Wiche 1998)	Cytosolic/Plasma membrane scaffolding
Actin, beta 14250401	97	R-K-C-F-COOH	-		
Beta actin 16359158	96				
Annexin A6 113962	150	G-G-E-D-COOH	-	Phospholipids (Gerke et al., 2001)	Cytosolic/Plasma membrane associated
Cytokeratin 8 181573	73				
Keratin type II cytosk 8 2506774	73	V-L-P-K-COOH	-	14-3-3, TNFR2 receptor, Caspase 9 (Coloumbe et al., 2004)	Cytosolic
Keratin 8 4504919	73				
Cytokeratin 10 gi:21961605	150			Many including: Plectin, Desmoplakin, BPAG1e, Plakophilin, Periplakin (Coloumbe et al., 2004)	Cytosolic
Cytokeratin 1 547749	150	G-P-R-Y-COOH	-		
α-filamen A 57284166	94			Actin, Glycoprotein-Ib α , Integrins ($\beta 1A$, $\beta 1D$, $\beta 2$, $\beta 3, \beta 7$), Dopamine D2, D3 receptor, Calveolin (van de Flier et al., 2001)	Plasma membrane associated
FLJ00343 protein 21748542	84	R-V-V-V-P-COOH	Class II		
Tubulin, beta polypeptide 57209813	172	E-E-E-A-COOH /	-	Many including: HSPs, MAPs (Gache et al., 2005)	Cytosolic
Tubulin, beta 4 15489150	120	E-E-V-A-COOH			
Tubulin beta 5 35959	102				
Tubulin, alpha 6 14389309	74				
Tubulin, alpha 6 variant 62897609	74	G-E-E-Y-COOH	-	Many including: HSPs, MAPs (Gache et al., 2005)	Cytosolic
plectin isoform 8 41322919	149				
plectin isoform 2/11 41322908/41322923	143	E-S-A-V-A-COOH	Class I/II	Actin, Tubulin, Desmoplakin, Hemidesmosomes, Vimentin, α -Spectrin (fodrin), Integrin, Cytokeratins (Wiche 1998)	Plasma membrane-associated
plectin isoform 3/10 41322908/41322914	141				
plectin isoform 6 41322916	139				
α-spectrin (α-fodrin) 55663122	106	S-L-F-V-N-COOH	Class II	Actin, protein 4.1, Plectin (Wiche 1998)	Plasma membrane associated
SPTAN1 protein 31565122	103				
Vimentin 5030431	150	D-D-L-E-COOH	None	Many including Actin, Tubulin, Integrins, Plectin, Desmin (Wiche 1998; Coloumbe et al., 2004)	Cytosolic

MALDI-TOF-MS analysis of tryptic peptides obtained from a reproducible band running at 105 kDa provisionally identified a member of the α -actinin family of proteins by peptide-mass fingerprinting (Figure 3.1B, C). MS identification of a non-muscle form of α -actinin after pull-down with the PDZ1 domain of ZO-1 was replicated using Cx36-transfected HeLa cells and wild-type glioma C6 cells (data not shown). Inspection of the C-terminus region of non-muscle α -actinin-1 and -4 indicated the presence of a class I PDZ domain interaction sequence (-S/T-X- ϕ ; E-S-D-L-COOH), suggesting that both of these proteins may be capable of associating with ZO-1 (Table 3.2).

Sequence Alignments of α -actinin-1 and -4 suggested a high degree of sequence homology (86%) (Table 3.2). Further tandem-MS analysis was used to confirm the identity of the 105 kDa protein before further molecular characterizations of the interaction with ZO-1. Prior to tandem-MS sequencing, peptides were chromatographically separated and deposited onto a MALDI target using SepDep vacuum chromatography as previously described in Part 2 (Chen et al., 2004). The application of the device resulted in greater than 20-fold increase in the signal-to-noise ratio of m/z 994.5 and 1150.7 ions (Fraction 12, Figure 3.1D) compared to the m/z signal-to-noise ratios of the same ions within spectra obtained after solid phase extraction (Figure 3.1B). This separation allowed the detection and sequencing of peptides, including pyroQQSNEHLR (Figure 3.1E) and pyroQQSNEHLRR (Figure 3.1F) that were specific to α -actinin-4. Neither of these peptides nor

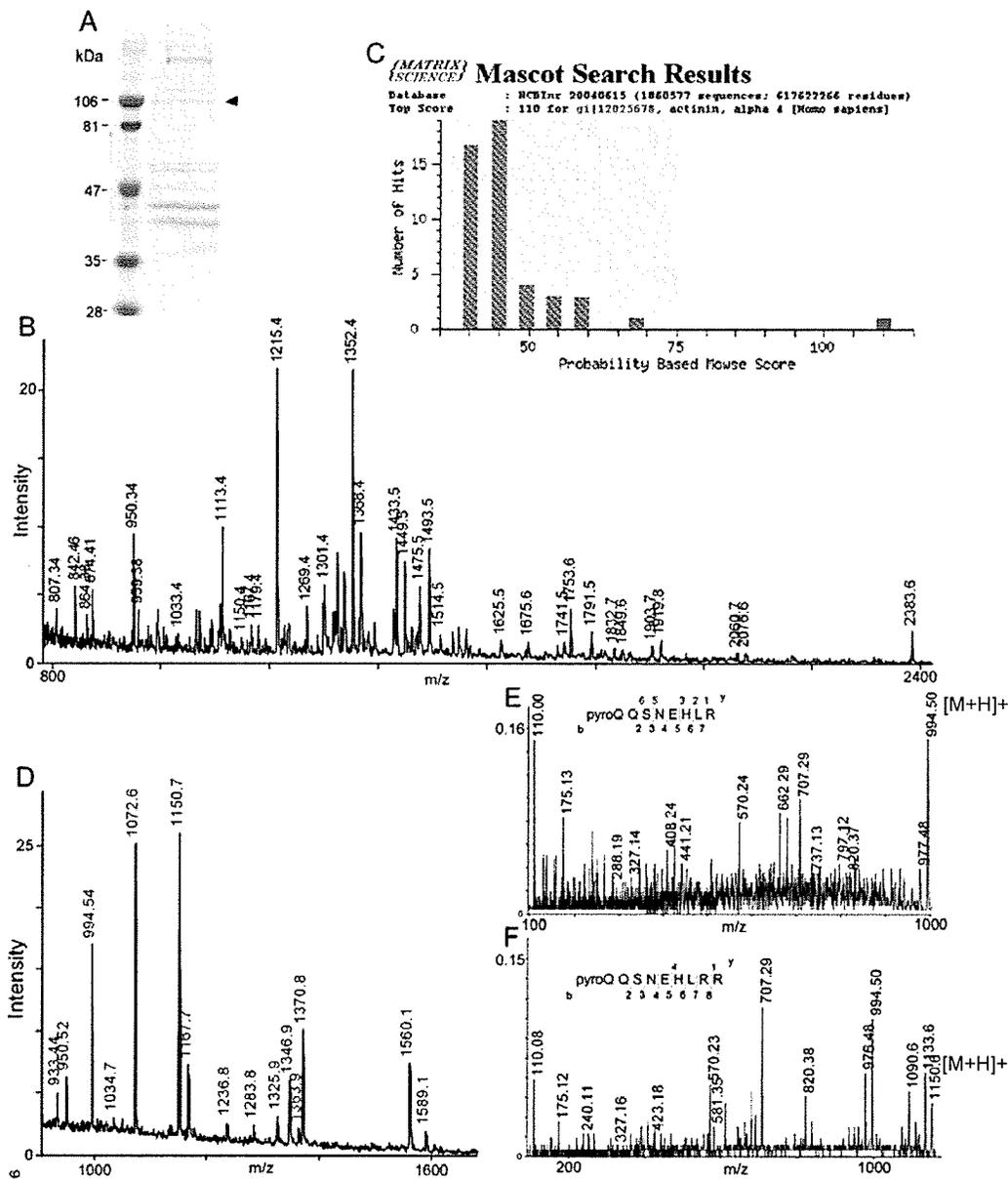


Figure 3.1: MS identification of α -actinin-4 following pull-down with GST fusion protein containing the PDZ1 domain of ZO-1. (A) Coomassie-stained gel of pull-down proteins, with arrowhead indicating the 105 kDa protein taken for MS analysis. (B) Spectrum and m/z ratios from tryptic fragments of the 105 kDa protein. (C) Database analysis of tryptic fragments showing a high probability that they were derived from nonmuscle α -actinin. (D) MS analysis of fraction 12 using SepDep liquid chromatography/MALDI. (E,F) Chromatographic separation and sequence identification of peptides having m/z at 994.5 (E) and m/z at 1150.6 (F), with the observed b and y fragmentation pattern (inset).

α-Actinin-1 (Sbjct) and α-Actinin-4 (Query) Sequence Alignment

Identities = 765/882 (86%), Positives = 821/882 (93%), Gaps = 0/882 (0%)

Query	11	DYMQPEEDWDRDLLDPAAWEKQQRKFTTAWCNSHLRKAGTQIENIEEDFRDGLKLMILLE	70
		DYM E+DWRDLLLDPAAWEKQQRKFTTAWCNSHLRKAGTQIENI+EDFRDGLKLMILLE	
Sbjct	30	DYMAQEDDWRDLLLDPAAWEKQQRKFTTAWCNSHLRKAGTQIENIEDFRDGLKLMILLE	89
Query	71	VISGERLAKPERGKMRVHKISNVNKALDFIASKGVKLVSIGAEIIVDGNVMTLGMIIWTI	130
		VISGERL KPERGKMRVHKI+NVNKALDFIASKGVKLVSIGAEIIVDGN KMTLGMIIWTI	
Sbjct	90	VISGERLPKPERGKMRVHKINNVNKALDFIASKGVKLVSIGAEIIVDGNAKMTLGMIIWTI	149
Query	131	ILRFAIQDISVEETSAKEGLLWCQRKTAPYKVNVIQNFHISWKDGLGFCALIHRRPEL	190
		ILRFAIQDISVEETSAKEGLLWCQRKTAPYKVN+QNFHISWKDGL F ALIHRRPEL	
Sbjct	150	ILRFAIQDISVEETSAKEGLLWCQRKTAPYKVNVIQNFHISWKDGLAFNALIHRRPEL	209
Query	191	IDYGKLRKDDPLTNLNTAFDVAEKYLDIPKMLDAEDIVGTARPDEKAIMTYVSSFYHAFS	250
		I+Y KLRKDDP+TNLN AF+VAEKYLDIPKMLDAEDIV TARPDEKAIMTYVSSFYHAFS	
Sbjct	210	IEYDKLRKDDPVTNLNNAFEVAEKYLDIPKMLDAEDIVNTARPDEKAIMTYVSSFYHAFS	269
Query	251	GAQKAETANRICKVLAVNQENEQLMEDYEKLASDLEWIRRTIPWLENRVPENTMHAHQ	310
		GAQKAETANRICKVLAVNQENE LMEDYEKLASDLEWIRRTIPWLE+RVP+ T+ MQ	
Sbjct	270	GAQKAETANRICKVLAVNQENEHLMEDYEKLASDLEWIRRTIPWLEDRVQKTIQEMQ	329
Query	311	QKLEDFRDYRRLHKPPKQVEKQCLEINFNTLQTKLRLSNRPAFMPSEGRMVSDDINNAWG	370
		QKLEDFRDYR+HKPPKQVEKQCLEINFNTLQTKLRLSNRPAFMPSEG+MVSDDINN W	
Sbjct	330	QKLEDFRDYRVRVHKPPKQVEKQCLEINFNTLQTKLRLSNRPAFMPSEGKMSVDDINNGWQH	389
Query	371	LEQVEKGYEEWLLNEIRRLERLDHLAEKFRQKASIEHAWTDGKEAMLKQDYETATLSBI	430
		LEQ EKGVEEWLLNEIRRLERLDHLAEKFRQKASIEHAWTDGKEAML+ +DYETATLS+I	
Sbjct	390	LEQAQKGYEEWLLNEIRRLERLDHLAEKFRQKASIEHAWTDGKEAMLKHRDYETATLSDI	449
Query	431	KALLKKHEAFESDLAAHQDRVEQIAAIAQELNELDYDPSVNRARQKICDQWDLGALT	490
		KAL++KHEAFESDLAAHQDRVEQIAAIAQELNELDYDPS +VN RQKICDQWDLG+LT	
Sbjct	450	KALIRKHEAFESDLAAHQDRVEQIAAIAQELNELDYDPSHNVNRARQKICDQWDLGSLT	509
Query	491	QKRREALERTEKLETTIDQLYLEYAKRAAFPNNWMEGAMEDLQDTFIVHTIEEIQGLTTA	550
		RREALE+TEK LE IDQL+LEYAKRAAFPNNWME AMEDLQD FIVHTIEEII+GL +A	
Sbjct	510	HSRREALERTEKLETTIDQLHLEYAKRAAFPNNWMEAMEDLQDMFIVHTIEEIEGLISA	569
Query	551	HEQFKATLPDADKERLAILGIHNEVSKIVQTYHVNMAAGTNPYTTITPQEINGKWDHVRQL	610
		H+QFK+TLPDAD+ER ALL IH E +I ++ H+ ++G+NPYTT+TPQ IN KW+ V+QL	
Sbjct	570	HDQFKSTLPDADREREAILAIHKEAQRIAESNHKLSGNSPYTTVTPQIINSKWEKVVQL	629
Query	611	VPRRDQALTEEHARQQHNERLRKQFGAQNVIQFWIQTKMEEIGRISIEMHGTLEDQLSH	670
		VP+RD AL EE ++QQ NE LR+QF +QANV+GFVIQTKMEEIGRISIEM+GTLEDQLSH	
Sbjct	630	VPKRDHALLEEQSKQSNELRRLRQFASQANVVGFWIQTKMEEIGRISIEMHGTLEDQLSH	689
Query	671	LRQYEKSIVNYKPKIDQLEGDRQLIQEALIFDNKHTNYTMEHIRVGWEQLLTTIARTINE	730
		L+QYE+SIV+YKP +D LE HQLIQEALIFDNKHTNYTMEHIRVGWEQLLTTIARTINE	
Sbjct	690	LKQYERSIVDYKPNLDLLEQQHQLIQEALIFDNKHTNYTMEHIRVGWEQLLTTIARTINE	749
Query	731	VENQILTRDAKGISQEQMNEFRASFNHFDKDRHSGTLGPEEFKACLISLGYDIGNDPQGEA	790
		VENQILTRDAKGISQEQM EFRASFNHFD+DH G LGPEEFKACLISLGYD+ ND QGEA	
Sbjct	750	VENQILTRDAKGISQEQMNEFRASFNHFDKDHGGALGPEEFKACLISLGYDVENDRQGEA	809
Query	791	EFARIMSVDPNRLGVVTFQAFIDFMSRETADTDADQVMASFKILAGDKNYITMDELRR	850
		EF RIMS+VDPN G+VTFQAFIDFMSRET DTDADQV+ASF+LAGDKN+IT +ELRR	
Sbjct	810	EFNRIMSLVDPNHSGLVTFQAFIDFMSRETTDADQVIASFVLAGDKNFITAEELRR	869
Query	851	ELFPDQAEYCIARMAPYTGPDVPGALDYMSFSTALYGESDL	892
		ELFPDQAEYCIARMAPY GPD+VPGALDY SFSTALYGESDL	
Sbjct	870	ELFPDQAEYCIARMAPYQGPDAVPGALDYKSFSTALYGESDL	911

Table 3.2: Sequence alignment between non-muscle α-actinin-1 (sbjct) and α-actinin-4 (query) demonstrating high sequence similarity (identity = 86%, center row) between α-actinin-1 and α-actinin-4. Alignment generated using NCBI protein-protein Blast (www.ncbi.nlm.nih.org).

their unpyrolyzed (+17 atomic mass units) forms were present in significant abundance for MS/MS prior to chromatographic separation.

3.4.2 Immunoblotting of α -actinin-4 from cultured cells and mouse tissues.

Lysates from cultured cells and homogenates from mouse tissues were taken for immunoblotting to confirm detection of α -actinin-4 with the anti- α -actinin4 antibodies employed. In Cx36-transfected HeLa cells, control and C6 glioma cells, anti- α -actinin-4 from ALEXIS detected a major band migrating at about 105 kDa, with very little difference in levels, as seen in blots loaded with equal protein (Figure 4.2A). A band at 105 kDa was also detected by this antibody in immunoblots of mouse liver and brain (Figure 3.2B) as well as heart (Figure 3.3A, lane 2). This band corresponds closely to the predicted molecular weight of α -actinin-4. Additional proteins of unknown identity exhibited much weaker reaction with this antibody, including a 38 kDa protein in liver and a 53 kDa protein that was faintly detected in most tissues examined. Similar results were obtained with anti- α -actinin-4 from Invitrogen/Zymed, as shown in an immunoblot of liver (Figure 3.2C), except that non-specific bands were less evident.

3.4.3 Co-IP of α -actinin-4 with ZO-1.

Molecular association of α -actinin-4 with ZO-1 was examined by co-IP of these proteins from homogenates of mouse tissues and lysates of cultured

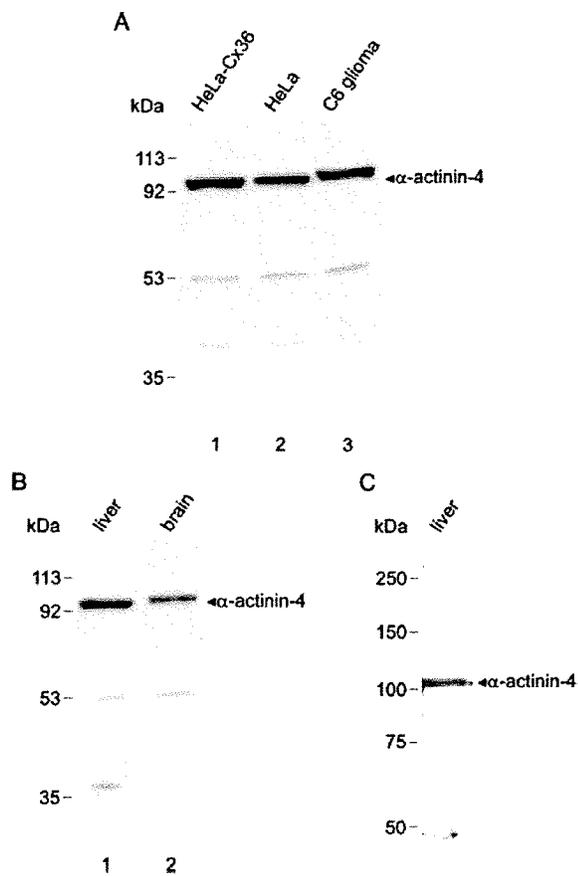


Figure 3.2: Immunoblots of cultured cells and mouse tissues probed with anti- α -actinin-4 antibodies. (A) Blots showing detection of α -actinin-4 in lysates of Cx36-transfected HeLa cells (lanes 1), control empty vector-transfected HeLa cells (lane 2), and C6 glioma cells (lane 3) with antibody from ALEXIS. (B, C) Blots showing detection of α -actinin-4 in homogenates of liver (B, lane 1) and brain (B, lane 2) with antibody from ALEXIS, and in homogenate of liver with antibody from Invitrogen/Zymed (C). In all blots, α -actinin-4 is seen migrating at about 105 kDa.

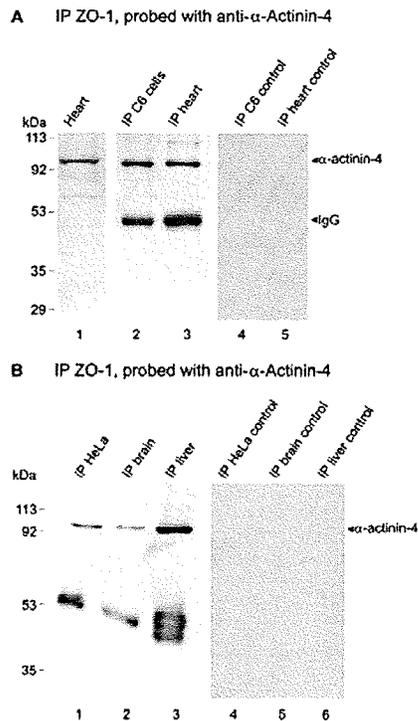


Figure 3.3: Co-IP of α -actinin-4 and ZO-1. Homogenates of cultured cells and mouse tissues were taken for IP with anti-ZO-1, and immunoblots of IP material were probed with anti- α -actinin-4. (A) Detection of α -actinin-4 in precipitates is seen after IP of ZO-1 from C6 glioma cells (lane 2) and heart (lane 3), with migration at 105 kDa corresponding to that of α -actinin-4 in whole homogenates of heart (lane 1). Control IP conducted with omission of anti-ZO-1 show an absence of α -actinin-4 using homogenates of C6 cells (lane 4) and heart (lane 5). (B) Detection of α -actinin-4 after IP of ZO-1 from HeLa cells (lane 1), brain (lane 2), and liver (lane 3). Control IP conducted with omission of anti-ZO-1 show an absence of α -actinin-4 in the same tissues (lanes 4-6). Bands in IP material migrating at about 50 kDa is IgG heavy chain.

cells. After IP of ZO-1 with Ab61-7300 from C6 glioma cells and heart, immunoblots of precipitates probed with anti- α -actinin-4 indicated the presence of α -actinin-4 in IP material (Figure 3.3A, lanes 2 and 3), with a migration profile corresponding to that of α -actinin-4 in whole homogenates of heart (Figure 3.3A, lane 1). Omission of anti-ZO-1 antibody (Figure 3.3A, lanes 4 and 5) served as negative controls. Similar results were obtained after IP of ZO-1 from HeLa cells, brain and liver, where IP material was consistently found to contain α -actinin-4 (Figure 3.3B, lanes 1, 2 and 3, respectively), with omission of IP antibody (Figure 3.3B, lanes 4, 5 and 6) serving as negative controls.

3.4.4 α -Actinin-4 binding to the PDZ1 domain of ZO-1.

In vitro pull-down assays involving each of the three PDZ domains of ZO-1 followed by immunoblotting of pull-down proteins were performed to confirm PDZ1 domain interaction with α -actinin-4, and to determine if the PDZ2 and PDZ3 domains of ZO-1 were also capable of associating with α -actinin-4. Each of the GST-PDZ1 to 3 domains were incubated with HeLa cell lysates, and pull-down material was probed with anti- α -actinin-4, with blots of whole lysates serving as a positive control for α -actinin-4 detection. As shown in Figure 3.4A, pull-down with the GST-PDZ1 domain resulted in detection of α -actinin-4 (Figure 3.4A, lane 2), corresponding to detection of α -actinin-4 in whole homogenates of HeLa cells (Figure 3.4A, lane 1), whereas pull-down

with GST-PDZ2 and GST-PDZ3 failed to reveal the presence of α -actinin-4 (Figure 3.4A, lanes 3 and 4). The immunoblot membrane in Figure 3.4A was stripped and reprobed with an anti-GST antibody to confirm equal loading of GST-fusion proteins at expected molecular weights of 40-42 kDa, and absence of the recombinant proteins in HeLa cell lysates serving as a negative control (Figure 3.4B, lane 1). As a positive control for the competence of GST-PDZ2 domain fusion protein to bind PDZ motifs, homogenates of HeLa cells stably transfected with Cx43 were used to pull-down Cx43 with this fusion protein. Consistent with previous reports of Cx43 interaction with the second PDZ domain of ZO-1 (Giepmans et al., 1998; Li et al., 2004^a), Cx43 was pulled-down with the PDZ2 domain, but not with the PDZ1 or PDZ3 domains of ZO-1 (Figure 3.4C). The absence of Cx43 interaction with PDZ1 serves as a negative control for the interaction of α -actinin-4 with this domain. Inclusion of a positive binding control for the PDZ3 domain was not possible because no known binding partner for this individual domain is currently known.

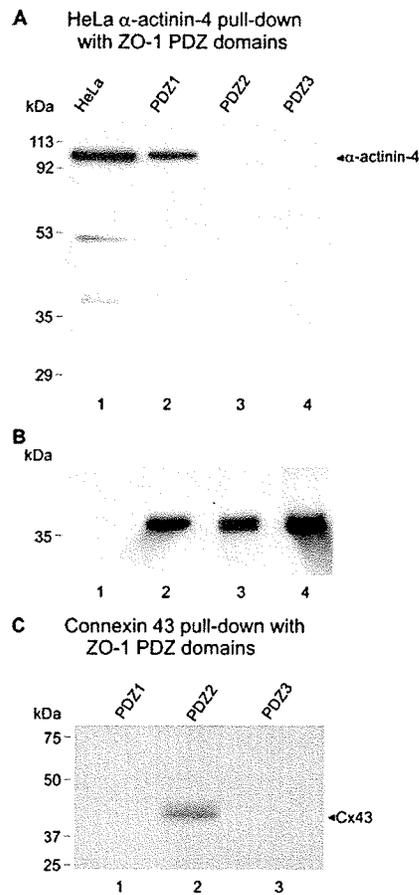


Figure 3.4: Pull-down analyses showing α -actinin-4 interaction with the PDZ1 domain of ZO-1. (A) Lysates of HeLa cells were incubated with GST fusion proteins containing the PDZ1, PDZ2, or PDZ3 domain of ZO-1. Immunoblots of pull-down material (lanes 2, 3, and 4) show binding of α -actinin-4 to the PDZ1 domain of ZO-1 (lanes 2) but not to the PDZ2 (lane 3) or the PDZ3 (lane 4) domains of ZO-1. Whole homogenates of HeLa cells were used as a positive control for detection of α -actinin-4 (lane 1). (B) The same membrane as in (A) was stripped and reprobed with anti-GST antibody to confirm equal interlane loading of GST fusion proteins, which are seen migrating at 40-42 kDa in pull-down material (lanes 2-4) and are absent in HeLa cell lysates (lane 1). (C) Positive control for the functionality of the second PDZ domain of ZO-1 shows pull-down of Cx43 with the PDZ2 domain (lane 2) and absence of Cx43 pull-down with the PDZ1 and PDZ3 domains (lanes 1 and 3).

3.4.5 Lack of α -actinin-1 co-IP with ZO-1.

Following immunoprecipitation of ZO-1 with Ab61-7300 from HeLa cells, MDCK cells and liver, IP material was probed with polyclonal anti- α -actinin-4 and blots were stripped and reprobed with monoclonal anti- α -actinin-1. There was detection of α -actinin-4 after IP (Figure 3.5A, lanes 1, 2, 3) and in cell lysates (Figure 3.5A, lanes 5 and 6), but not α -actinin-1 (Figure 3.5B, lanes 1, 2, 3) in IP material. Conversely, following immunoprecipitation of ZO-1 with Ab61-7300 from the same tissues, IP material was probed with monoclonal anti- α -actinin-1 (Figure 3.5C) and blots were stripped and reprobed with polyclonal anti- α -actinin-4. The blots showed lack of α -actinin-1 detection (Figure 3.5C, lanes 1, 2, 3), but detection of α -actinin-4 (Figure 3.5B, lanes 1, 2, 3). Positive controls for detection of α -actinin-1 in homogenates of MDCK and HeLa cells are shown in Figure 3.5B.

3.4.6 Immunofluorescence co-localization of α -actinin-4 with ZO-1.

The distribution and co-localization relationships of α -actinin-4 and ZO-1 were examined by double immunofluorescence labeling of these proteins in cultures of wild-type HeLa cells, Cx36-transfected HeLa cells and C6 glioma cells. In wild-type HeLa cells, ZO-1 appeared as punctate or short strands of labeling along plasma membranes at points of cell-cell contacts, and only faint labeling was evident in the cytoplasmic or nuclear compartments (Figure 3.6A1). Immunolabeling of α -actinin-4 was also present at cell-cell contacts (Figure 3.6A2), where it was often co-localized with ZO-1 (Figure 3.6A3).

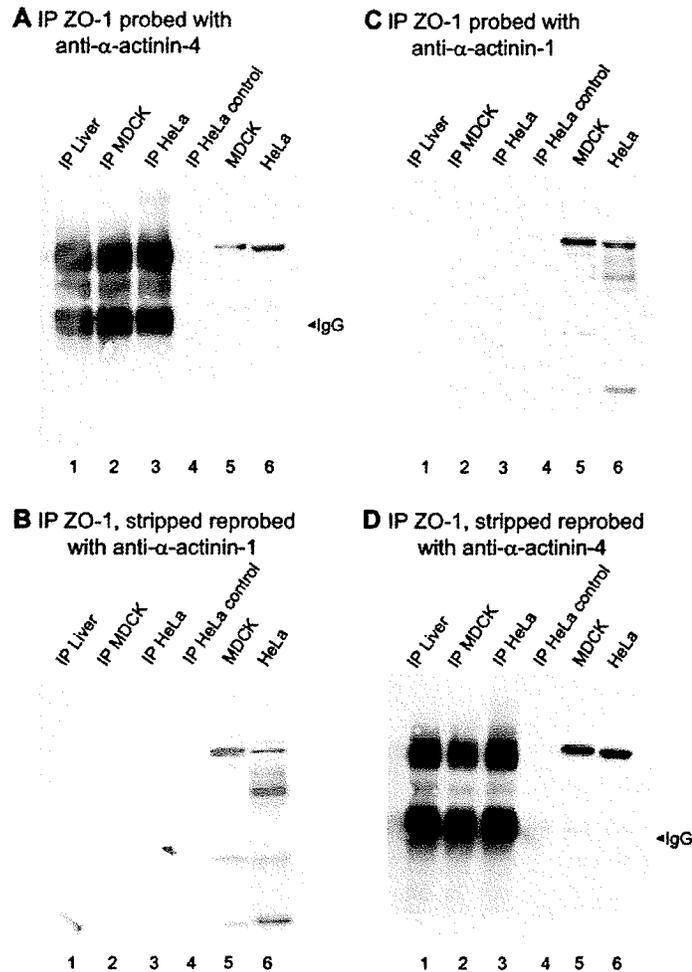


Figure 3.5: Lack of α -actinin-1 association with ZO-1. (A) Immunoblot probed with anti- α -actinin-4 after IP of ZO-1 from homogenates of liver (lane 1), MDCK cells (lane 2), and HeLa cells (lane 3), with antibody omission from HeLa cells serving as a negative control (lane 4) and probing of MDCK and HeLa cells serving as positive controls for α -actinin-4 detection (lanes 5 and 6). (B) Blot in A was stripped and reprobred with antibody against α -actinin-1 and shows absence of α -actinin-1 in IP material but shows its presence in MDCK (lane 5) and HeLa (lane 6) cells. (C) Immunoblot probed with anti- α -actinin-1 after IP of ZO-1 from homogenates of liver (lane 1), MDCK cells (lane 2), and HeLa cells (lane 3), with antibody omission from HeLa cells serving as a negative control (lane 4) and probing of MDCK and HeLa cells serving as positive controls for α -actinin-1 detection (lanes 5 and 6). (D) The blot in C was stripped and reprobred with antibody against α -actinin-4, showing as positive controls the presence of α -actinin-4 in IP material and its presence in MDCK (lane 5) and HeLa (lane 6) cells caused a reduction in cytoplasmic levels of α -actinin-4 and concomitantly a large accumulation of α -actinin-4 at plasma membranes (Figure 3.6B2).

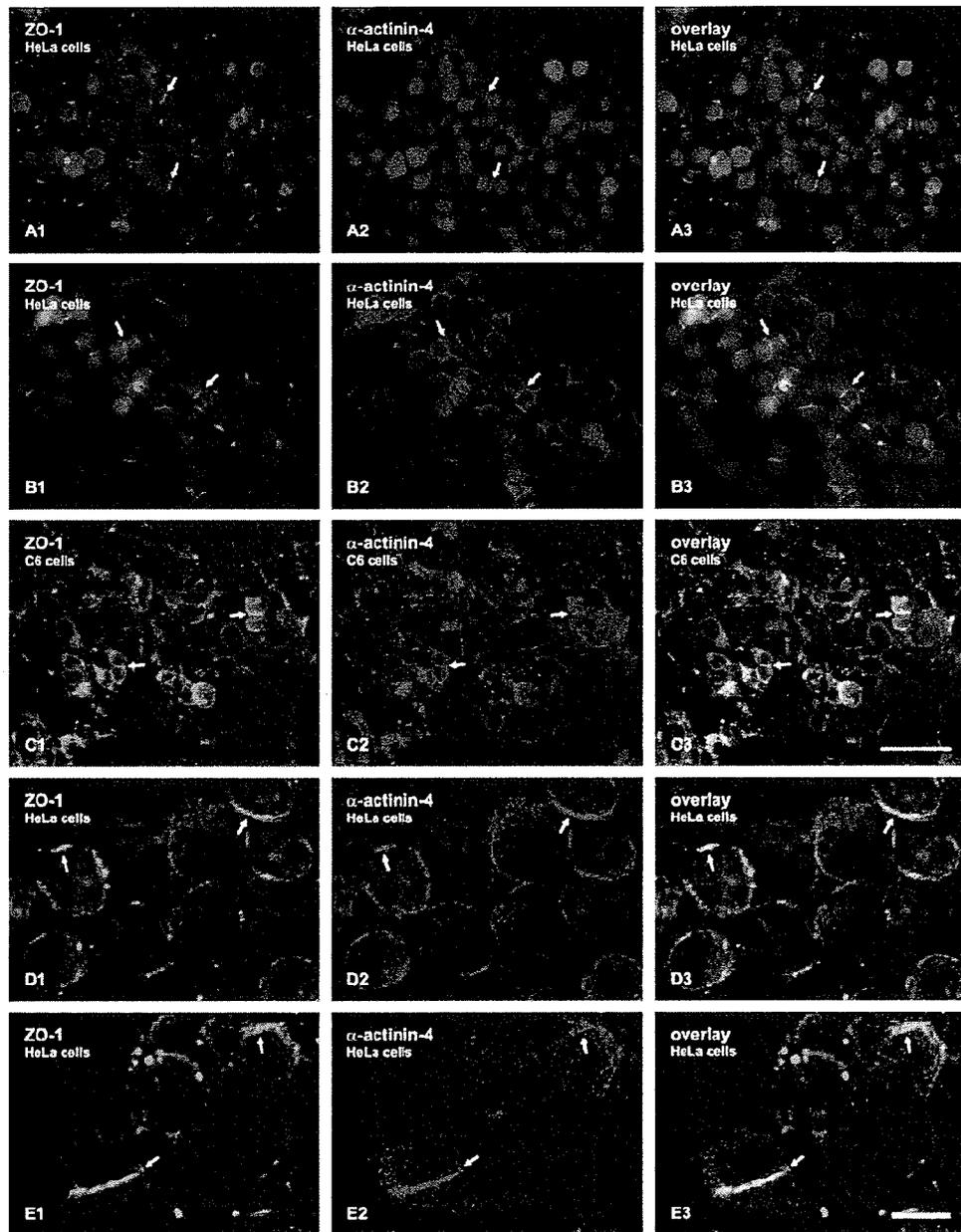


Figure 3.6: Immunofluorescence labeling of ZO-1 and α -actinin-4 in cultured HeLa cells and C6 glioma cells. (A) Double immunofluorescence of the same field in cultures of wild-type HeLa cells showing ZO-1 localized as puncta and as strands of labeling at points of cell-cell contact (A1, arrows), and α -actinin-4 within cytoplasm and points of cell contacts (A2, arrows), with partial α -actinin-4/ZO-1 colocalization at cell contacts, as seen in image overlay (A3, arrows). (B) Double immunofluorescence of the same field in cultures of

Cx36-transfected HeLa cells showing more prominent localization of ZO-1 at cell contacts (B1, arrows) and a greater abundance of α -actinin-4 at cell appositions with a reduction in cytoplasmic labeling (B2, arrows), and a concomitant increase in ZO-1/ α -actinin-4 colocalization (B3, arrows). (C) Double immunofluorescence of the same field in cultures of C6 glioma cells showing ZO-1 (C1) and α -actinin-4 (C2) at sites of cell-cell appositions (arrows) and α -actinin-4/ZO-1 colocalization at these sites (C3, arrows). (D,E) Higher magnification laser scanning confocal images showing double labeling for α -actinin-4 and ZO-1 in control HeLa cells (D) and in Cx36-transfected HeLa cells (E), with arrows indicating α -actinin/ZO-1 colocalization. Scale bars: A-C, 50 μ m; D,E, 10 μ m and 5C, lanes 5 and 6. All four blots included a negative control (lane 4) showing absence of bands after IP from HeLa cells with omission of ZO-1 primary antibody.

Immunostaining of α -actinin-4 in these cells was also widely distributed intracellularly, which tended partially to obscure labeling along plasma membranes. In Cx36-transfected HeLa cells, ZO-1 appeared and extensively co-localized with α -actinin-4 (Figure 3.6B3). At present, we have no explanation for this cellular redistribution of α -actinin-4 in cells overexpressing Cx36, but which do not normally express this or any other connexin. Higher magnification laser scanning confocal images of ZO-1/ α -actinin-4 association in nontransfected HeLa cells and Cx36-transfected HeLa cells are shown in Figures 3.6D and 3.6E respectively. In C6 glioma cells, immunolabeling of both ZO-1 (Figure 3.6C1) and α -actinin-4 (Figure 3.6C2) was observed weakly in the cytoplasm, but more prominently along plasma membranes, where the two proteins were often co-localized (Figure 3.6C3).

Association of α -actinin-4 with ZO-1 was examined by double immunofluorescence in various tissues of adult mouse. In cerebral cortex and heart, labeling of ZO-1 was typically seen concentrated along blood vessels, specifically as linear strands of immunofluorescence (Figure 3.7A1 and 3.7B1). In both tissues, α -actinin-4 had a similar distribution on vessels (Figs 3.7A2 and 3.7B2) and was largely co-localized with ZO-1 along linear strands of labeling (Figure 3.7A3 and 3.7B3). In liver, intense labeling for ZO-1 was distributed at the periphery of cells, outlining individual hepatocytes (Figure 3.7C1), and α -actinin-4 had a similar distribution (Figure 3.7C2), often overlapping that of ZO-1 (Figure 3.7C3). In lung, labeling of ZO-1 occurred at the plasma membranes in cells of respiratory bronchioles (Figure 3.6D1) in a pattern matching that of α -actinin-4 (Figure 3.6D2), which gave rise to extensive co-localization of the two proteins (Figure 3.6D3).

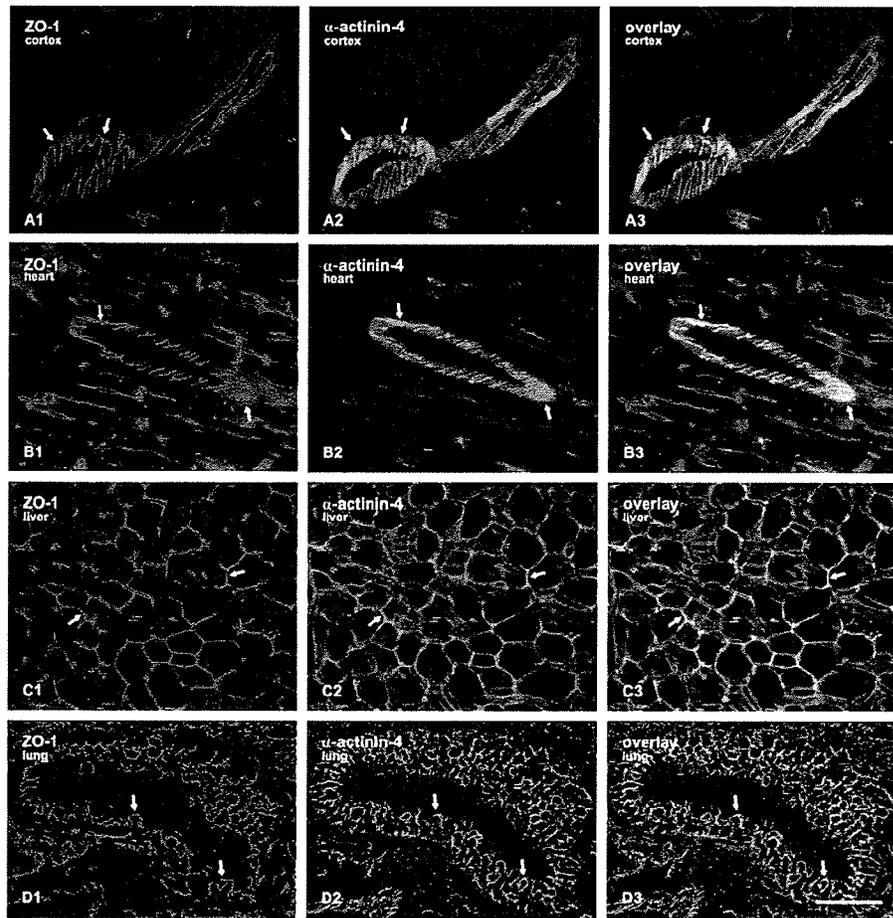


Figure 3.7: Double immunofluorescence labeling of ZO-1 and α -actinin-4 in tissues of adult mouse. (A,B) The same field in cerebral cortex (A) and heart (B) showing strands (arrows) of labeling for ZO-1 (A1,B1) and α -actinin-4 (A2,B2) along blood vessels, and ZO-1/ α -actinin-4 colocalization along these strands, as seen by yellow in image overlay (A3,B3, arrows). (C) The same field in liver showing labeling of ZO-1 (C1) and α -actinin-4 (C2) around hepatocytes, and a high degree of ZO-1/ α -actinin-4 colocalization (arrows), as seen by yellow in image overlay (C3). (D) The same field showing labeling of ZO-1 (D1) and α -actinin-4 (D2) in a respiratory bronchiole of lung, with substantial colocalization of labeling (D3). Scale bars: 50 μ m.

3.5 Discussion and conclusions

We have demonstrated that functional proteomics is an effective means of characterizing interaction networks of PDZ domain-containing proteins and complexes. By using the first of three PDZ domains in ZO-1 to pull-down proteins for MALDI-MS identification, a number of potential ZO-1 interacting candidates were found based on c-terminal sequence evaluation that indicated the presence of PDZ domain consensus binding motifs. Among these candidates, we established that α -actinin-4 specifically interacts with the PDZ1 domain, but not the PDZ2 or PDZ3 domains of ZO-1. This interaction was supported by co-IP of these proteins and their co-localization in cultured cells and various tissues *in vivo*. In contrast, α -actinin-1 failed to show an association with ZO-1, suggesting that α -actinin-1 and α -actinin-4 are functionally distinct members of the non-muscle α -actinins. Using MS, we identified several other components within ZO-1 membrane-associated protein complexes, including F-actin, plectin, tubulin, filamin A, cytokeratins, α -spectrin/fodrin and vimentin, as summarized in Table 3.1. Among these are proteins that contain putative PDZ binding motifs (Table 3.1) and may bind PDZ domains, or that may originate from higher order protein complexes involving ZO-1. Several of these were previously found either to have distributions resembling ZO-1 or to be co-localized with ZO-1 at the leading edge of lamellipodia in fibroblasts, keratocytes and myofibroblasts, which suggests that they are important components of the ZO-1 scaffolding network

(Torrado et al., 2004; Wiche 1998; Taliana et al., 2005), but further detailed molecular studies are required to determine their relationship with ZO-1.

3.5.1 MS-based identification of protein interactions.

The use of recombinant “bait” proteins to capture protein binding partners for identification by MS is powerful, but still suffers from several drawbacks, particularly in studies of protein interaction networks. The first of these is the potential for obtaining false positive interactions that occur when protein “baits” are exposed to non-physiological environments created by tissue disruption, resulting in complex mixtures prone to formation of non-specific protein aggregates. This complexity was reduced by our use of a discrete region of the ZO-1 sequence as bait, namely its PDZ1 interaction domain. PDZ modules bind to a single, short, c-terminus peptide sequence within a target protein and thus require far less discriminating power than protein interaction domains utilizing internal sequences, which require full analysis of the entire protein sequence with considerations of protein structure and surface accessibility. Thus, from a large array of MS-identified proteins in the PDZ1 pull-down, our screening for PDZ interacting candidates was based on the evaluation of the last four c-terminal amino acids, thereby simplifying evaluation of potential protein binding motifs based on linear primary sequence. These unique features of PDZ domains and their mode of association offer a powerful yet simplified constraint for the identification of first order PDZ-protein interactions.

A second means of reducing false positives was elution of PDZ1 domain-associated proteins and complexes with a competitor peptide corresponding to the c-terminus region of Cx36, which contains a class I (S/T-X- Φ) and II (Φ -X- Φ) PDZ binding sequence within the last four amino acids (S-A-Y-V-COOH). It was previously demonstrated that this peptide inhibited Cx36 interaction with ZO-1 (Li et al 2004^a; Li et al., 2004^b), indicating that it could be used to selectively release PDZ1-associated proteins for proteomic identification. Elution with this peptide resulted in a dramatic reduction of non-specific background, and reduced the levels of contamination by GST-fusion protein in SDS-PAGE gels.

Our application of a novel vacuum enabled micro-liquid chromatography (SepDep) device for MALDI-MS analyses (see Part 2, Chen et al., 2004) further addressed the problem of protein complexity. The device allows simple and effective vacuum-based LC separation that simultaneously deposits LC derived fractions onto a MALDI surface. Although solid phase extraction, such as ZipTips™, is effective in removing ionic and polar impurities, interfering (high abundance and readily ionized) peptides also are often concentrated and limit proteomic analysis. In addition, high levels of homologous sequences within distinct proteins (such as the α -actinins), together with ion suppression, reduce the number and relative abundance of sequence specific peptides for fragmentation analysis. Application of the

device allowed the identification of α -actinin-4 specific peptides within the sample in a rapid and cost effective manner.

3.5.2 ZO-1/ α -actinin-4 interaction.

The finding that α -actinin-4 has PDZ domain interaction capability is consistent with a previous report showing that the PDZ binding class I sequence at the C-terminus of α -actinin-4 (E-S-D-L-COOH) interacts with the fifth PDZ domain of the tight junction-associated protein MAGI-1 (Patrie *et al.*, 2002). Our demonstration of α -actinin-4 co-localization with another major component of tight junctions, namely ZO-1, provides a second means for targeting α -actinin-4 to these structures. More generally, it is well established that α -actinin-4 is highly concentrated at a variety of other cell-cell junctions, cell-matrix contacts and plasma membrane receptor systems, some of which also harbor ZO-1. Thus, our results have revealed a protein-protein interaction mode whereby ZO-1 may be targeted to these structures for the purpose of intracellular signaling and/or scaffolding between plasma membrane components and the actin-based cytoskeleton, with which both α -actinin-4 and ZO-1 are also intimately associated. It is noteworthy that along with its three PDZ domains, ZO-1 also contains a 220 amino acid carboxy-terminal region termed the actin-binding domain that associates with F-actin and that was thought to be largely responsible for localization of ZO-1 to the cortical cytoskeleton (Itoh *et al.*, 1997; Fanning *et al.*, 1998; Wittchen *et al.*, 1999). However, ZO-1 constructs lacking the actin-binding domain were still

able to localize to plasma membranes as well as to cell-cell contacts (Fanning et al., 1998), indicating that its recruitment to the plasma membrane and the cytoskeleton are mediated by additional domains, including its PDZ1 domain interaction with α -actinin-4, as shown by the present results.

3.5.3 Functional implications.

The α -actinins are a family of closely related proteins that arise from four separate genes with each molecule containing an N-terminal actin-binding domain, four α -helical spectrin-like repeats and a c-terminal calmodulin-like domain (Djinovic-Carugo et al., 2002; Otey et al., 2004). α -Actinins-2 and -3 are highly expressed in muscle, whereas α -actinins-1 and -4 are expressed in many cell types, where they are present at the cytoplasmic face of plasma membranes (Otey et al., 2004; Beggs et al., 1992). α -Actinin-1 is primarily associated with focal adhesions and the F-actin stress fibers (Belkin et al., 1987; Wachsstock et al., 1987; Otey et al., 1990; Parast et al., 2000; Pomies et al., 1997). α -Actinins interact with α -catenin, β -catenin and protein ADIP at adherens junctions (Knudsen et al., 1995; Hayashida et al., 2005; Asada et al., 2003), with vinculin, β 1, β 3 integrins, palladin and protein CRP at focal adhesions (Belkin et al., 1987; Wachsstock et al., 1987, Wachsstock et al., 1987; Otey et al., 1990; Parast et al., 2000; Pomies et al., 1997), and with protein BP180 at hemidesmosomes (Gonzalez et al., 2001). In some cells, they also interact with the intercellular domains of ICAM-1, L-selectin and β 2 integrins (Carpen et al., 1992; Heiska et al., 1996;

Pavalko et al., 1993; Pavalko et al., 1995). Given its multi-functional roles in relation to transmembrane proteins and cytoskeletal networks, α -actinin-4 has received considerable attention as a target for regulation by signaling pathways, particularly in the areas of metastasis and cancer. Assembly and disassembly of α -actinin-4 and associated proteins at adherens junctions was shown to play an important role in either stabilizing cell adhesion and suppressing cell motility or promoting loss of adherens junctions leading to increased rates of metastasis (Gluck et al., 1993; Gluck et al., 1994; Guvakova et al., 2002). Further, up-regulation of α -actinin-4 was correlated with increased cell motility, invasive growth and lymph node metastasis in colorectal cancer (Honda et al., 2005), and was a prognostic predictor in non-small cell lung cancers (Yamagata et al., 2003). Greatly elevated levels of cytoplasmic α -actinin-4 were found in invasive phenotypes of breast cancer with high metastatic potential and poor survival rates (Honda et al., 1998).

The present findings provide new avenues for exploring mechanisms whereby up-regulation of cytoplasmic α -actinin-4 contribute to mitogenic signaling. One such avenue is related to the reported interaction of ZO-1 with the Y-box transcription factor ZONAB (zonula occludens-1 associated nucleic acid binding protein), which was localized both at tight junctions and in cell nuclei (Balda et al., 2000; Balda et al., 2003). In the nucleus, ZONAB binds to the promoter region of the proto-oncogene ErbB-2, acting as a transcriptional repressor of this gene in a cell density-dependent manner. ErbB-2

(HER2/neu) is a receptor tyrosine kinase that is overexpressed in 25-30% of diagnosed breast cancers (Slamon et al., 1987; Slamon et al., 1989; Emens et al., 2005). Studies have shown accumulation of cytoplasmic α -actinin-4 in various types of tumor cells (Honda et al 1998; Yamagata et al., 2003; Honda et al., 2005). Thus, overexpression of α -actinin-4 may trap ZONAB in the cytoplasm via α -actinin-4/ZO-1/ZONAB association and relieve ZONAB repression of ErbB-2 gene, thereby promoting ErbB-2-mediated tumorigenesis. Alternatively, α -actinin-4 has been shown to undergo nuclear translocation, indicating its involvement in signaling platforms beyond the plasma membrane. Such nuclear translocation was correlated with less aggressive breast carcinomas (Honda et al., 1998). Although α -actinin-4 does not contain nuclear import sequences (Honda et al., 2005), several such sequences occur in ZO-1 (Islas et al., 2002; Balda et al., 2000). Thus, our findings raise the possibility of α -actinin-4/ZO-1 co-transport to the nucleus, where α -actinin-4 appears to exert regulatory control of growth and/or metastatic potential. Interestingly, although α -actinin-4 has been observed within the nucleus, α -actinin-1 which fails to interact with ZO-1 (as presented here), has demonstrated no such nuclear translocation ability (Honda et al., 1998).

3.6 References

- Asada, M., Irie, K., Morimoto, K., Yamada, A., Ikeda, W., Takeuchi, M., Takai, Y. (2003) ADIP, a novel Afadin- and alpha-actinin-binding protein localized at cell-cell adherens junctions. *J. Biol. Chem.* 278, 4103-4111.
- Balda, M. S., Garrett, M. D., Matter, K. (2003) The ZO-1-associated Y-box factor ZONAB regulates epithelial cell proliferation and cell density. *J. Cell Biol.* 160, 423-432.
- Balda, M. S., Matter, K. (2000) The tight junction protein ZO-1 and an interacting transcription factor regulate ErbB-2 expression. *EMBO J.* 19, 2024-2033.
- Beggs, A. H., Byers, T. J., Knoll, J. H., Boyce, F. M., Bruns, G. A., Kunkel, L. M. (1992) Cloning and characterization of two human skeletal muscle alpha-actinin genes located on chromosomes 1 and 11. *J. Biol. Chem.* 276, 9281-9288.
- Belkin, A. M., Kotliansky, V. E. (1987) Interaction of iodinated vinculin, metavinculin and alpha-actinin with cytoskeletal proteins. *FEBS Lett.* 220, 291-294.
- Carpen, O., Pallai, P., Staunton, D. E., Springer, T. A. (1992) Association of intercellular adhesion molecule-1 (ICAM-1) with actin-containing cytoskeleton and alpha-actinin. *J. Cell Biol.* 118, 1223-1234.
- Chen, V.C., Cheng, K., Ens, W., Standing, K. G., Nagy, J. I., Perreault, H. (2004) Device for the reversed-phase separation and on-target deposition of peptides incorporating a hydrophobic sample barrier for matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 76, 1189-1196.
- Coulombe, P. A., Wong, P. (2004) Cytoplasmic intermediate filaments revealed as dynamic and multipurpose scaffolds. *Nat. Cell Bio.* 6, 699-706.
- Djinovic-Carugo, K., Gautel, M., Ylanne, J., Young, P. (2002) The spectrin repeat: a structural platform for cytoskeletal protein assemblies. *FEBS Lett.* 513, 119-123.
- Emens, L. A. (2005) Trastuzumab: targeted therapy for the management of HER-2/neu-overexpressing metastatic breast cancer. *Am. J. Ther.* 12, 243-253.

- Fanning, A. S., Jameson, B. J., Jesaitis, L. A., Anderson, J. M. (1998) The tight junction protein ZO-1 establishes a link between the transmembrane protein occludin and the actin cytoskeleton. *J. Biol. Chem.* 273, 29745-29753.
- Fanning, A. S., Ma, T. Y., Anderson, J. M. (2002) Isolation and functional characterization of the actin binding region in the tight junction protein ZO-1. *FASEB J.* 16, 1835-1837.
- Gache, V., Louwagie, M., Garin, J., Caudron, N., Lafanechere, L., Valiron, O. (2005) Identification of proteins binding the native tubulin dimer. *Res. Commun.* 327, 35-42.
- Gerke, V., Moss, S. E. (2001) Annexins: from structure to functions. *Physiol Rev.* 82, 331-371.
- Giepmans, B. N., Moolenaar, W. H. (1998). The gap junction protein connexin43 interacts with the second PDZ domain of the zona occludens-1 protein. *Curr. Biol.* 8, 931-934.
- Gluck, U., Ben-Ze'ev, A. (1994). Modulation of alpha-actinin levels affects cell motility and confers tumorigenicity on 3T3 cells. *J. Cell Sci.* 107, 1773-1782.
- Gluck, U., Kwiatkowski, D. J., Ben-Ze'ev, A. (1993). Suppression of tumorigenicity in simian virus 40-transformed 3T3 cells transfected with alpha-actinin cDNA. *Proc. Natl. Acad. Sci. USA.* 90, 383-387.
- Gonzalez, A. M., Otey, C., Edlund, M., Jones, J. C. (2001) Interactions of a hemidesmosome component and actinin family members. *J. Cell Sci.* 114, 4197-4206.
- Gottardi, C. J., Arpin, M., Fanning, A. S., Louvard, D. (1996) The junction-associated protein, zonula occludens-1, localizes to the nucleus before the maturation and during the remodeling of cell-cell contacts. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10779-10784.
- Guvakova, M. A., Adams, J. C., Boettiger, D. (2002) Functional role of alpha-actinin, PI 3-kinase and MEK1/2 in insulin-like growth factor I receptor kinase regulated motility of human breast carcinoma cells. *J. Cell Sci.* 115, 4149-4165.
- Hayashida, Y., Honda, K., Idogawa, M., Ino, Y., Ono, M., Tsuchida, A., Aoki, T., Hirohashi, S., Yamada. (2005) T. E-cadherin regulates the association between beta-catenin and actinin-4. *Cancer Res.* 65, 8836-8845.

- Heiska, L., Kantor, C., Parr, T., Critchley, D. R., Vilja, P., Gahmberg, C. G., Carpen, O. (1996) Binding of the cytoplasmic domain of intercellular adhesion molecule-2 (ICAM-2) to alpha-actinin. *J. Biol. Chem.* 271, 26214-26219.
- Honda, K., Yamada, T., Endo, R., Ino, Y., Gotoh, M., Tsuda, H., Yamada, Y., Chiba, H., Hirohashi, S. (1998) Actinin-4, a novel actin-bundling protein associated with cell motility and cancer invasion. *J. Cell Biol.* 140, 1383-1393.
- Honda, K., Yamada, T., Hayashida, Y., Idogawa, M., Sato, S., Hasegawa, F., Ino, Y., Ono, M., Hirohashi, S. (2005) Actinin-4 increases cell motility and promotes lymph node metastasis of colorectal cancer. *Gastroenterology.* 128, 51-62.
- Inagaki, M., Irie, K., Deguchi-Tawarada, M., Ikeda, W., Ohtsuka, T., Takeuchi, M., Takai, Y. (2003) Nectin-dependent localization of ZO-1 at puncta adherentia junctions between the mossy fiber terminals and the dendrites of the pyramidal cells in the CA3 area of adult mouse hippocampus. *J. Comp. Neurol.* 460, 514-524.
- Islas, S., Vega, J., Ponce, L., Gonzalez-Mariscal, L. (2002) Nuclear localization of the tight junction protein ZO-2 in epithelial cells. *Exp. Cell Res.* 274, 138-148.
- Itoh, M., Nagafuchi, A., Yonemura, S., Kitani-Yasuda, T., Tsukita, S., Tsukita, S. (1993) The 220-kD protein colocalizing with cadherins in non-epithelial cells is identical to ZO-1, a tight junction-associated protein in epithelial cells: cDNA cloning and immunoelectron microscopy. *J. Cell Biol.* 121, 491-502.
- Itoh, M., Furuse, M., Morita, K., Kubota, K., Saitou, M., Tsukita, S. (1999) Direct binding of three tight junction-associated MAGUKs, ZO-1, ZO-2, and ZO-3, with the COOH termini of claudins. *J. Cell Biol.* 147, 1351-1363.
- Itoh, M., Nagafuchi, A., Moroi, S., Tsukita, S. (1997) Involvement of ZO-1 in cadherin-based cell adhesion through its direct binding to catenin and actin filaments. *J. Cell Biol.* 138, 181-192.
- Kausalya, P. J., Reichert, M., Hunziker, W. (2001) Connexin45 directly binds to ZO-1 and localizes to the tight junction region in epithelial MDCK cells. *FEBS Lett.* 505, 92-96.
- Knudsen, K. A., Soler, A. P., Johnson, K. R., Wheelock, M. J. (1995) Interaction of alpha-actinin with the cadherin/catenin cell-cell adhesion complex via alpha-catenin. *J. Cell Biol.* 130, 67-77.

- Laing, J. G., Manley-Markowski, R. N., Koval, M., Civitelli, R., Steinberg, T. H. (2001) Connexin45 interacts with zonula occludens-1 in osteoblastic cells. *Cell. Commun. Adhes.* 8, 209-212.
- Li, X., Olson, C., Lu, S., Kamasawa, N., Yasumura, T., Rash, J. E., Nagy, J. I. (2004) Neuronal connexin36 association with zonula occludens-1 protein (ZO-1) in mouse brain and interaction with the first PDZ domain of ZO-1. *Eur. J. Neurosci.* 19, 2132-2146.
- Li, X., Olson, C., Lu, S., Nagy, J. I. (2004) Association of connexin-36 with zonula occludens-1 in HeLa cells, betaTC-3 cells, pancreas, and adrenal gland. *Histochem. Cell. Biol.* 122, 485-498.
- Loboda, A., Krutchinsky, A. N., Bromirski, M., Ens, W., Standing, K. G. (2000) A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance. *Rapid Commun. Mass Spectrom.* 14, 1047-1057.
- Mitic, L. L., Schneeberger, E. E., Fanning, A. S., Anderson, J. M. (1999) Connexin-occludin chimeras containing the ZO-binding domain of occludin localize at MDCK tight junctions and NRK cell contacts. *J. Cell Biol.* 146, 683-693.
- Nielsen, P. A., Baruch, A., Shestopalov, V. I., Giepmans, B. N., Dunia, I., Benedetti, E. L., Kumar, N. M. (2003) Lens connexins alpha3Cx46 and alpha8Cx50 interact with zonula occludens protein-1 (ZO-1). *Mol. Biol. Cell.* 14, 2470-2481.
- Nielsen, P. A., Beahm, D. L., Giepmans, B. N., Baruch, A., Hall, J. E., Kumar, N. M. (2002) Molecular cloning, functional expression, and tissue distribution of a novel human gap junction-forming protein, connexin-31.9. Interaction with zona occludens protein-1. *J. Biol. Chem.* 277, 38272-38283.
- Otey, C. A., Carpen, O. (2004) Alpha-actinin revisited: a fresh look at an old player. *Cell Motil. Cytoskeleton.* 58, 104-111.
- Otey, C. A., Pavalko, F. M., Burridge, K. (1990) An interaction between alpha-actinin and the beta 1 integrin subunit in vitro. *J. Cell Biol.* 111, 721-729.
- Parast, M. M., Otey, C. A. (2000) Characterization of palladin, a novel protein localized to stress fibers and cell adhesions. *J. Cell Biol.* 150, 643-656.

- Patrie, K. M., Drescher, A. J., Welihinda, A., Mundel, P., Margolis, B. (2002) Interaction of two actin-binding proteins, synaptopodin and alpha-actinin-4, with the tight junction protein MAGI-1. *J. Biol. Chem.* 277, 30183-30190.
- Pavalko, F. M., LaRoche, S. M. (1993) Activation of human neutrophils induces an interaction between the integrin beta 2-subunit (CD18) and the actin binding protein alpha-actinin. *J. Immunol.* 151, 3795-3807.
- Pavalko, F. M., Walker, D. M., Graham, L., Goheen, M., Doerschuk, C. M., Kansas, G. S. (1995) The cytoplasmic domain of L-selectin interacts with cytoskeletal proteins via alpha-actinin: receptor positioning in microvilli does not require interaction with alpha-actinin. *J. Cell. Biol.* 129, 1155-1164.
- Pomies, P., Louis, H. A., Beckerle, M. C. (1997) CRP1, a LIM domain protein implicated in muscle differentiation, interacts with alpha-actinin. *J. Cell Biol.* 139, 157-168.
- Rajasekaran, A. K., Hojo, M., Huima, T., Rodriguez-Boulan, E. (1996) Catenins and zonula occludens-1 form a complex during early stages in the assembly of tight junctions. *J. Cell Biol.* 132, 451-463.
- Slamon, D. J. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science.* 235, 177-182.
- Slamon, D. J. (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science.* 244, 707-712.
- Songyang, Z., Fanning, A. S., Fu, C., Xu, J., Marfatia, S. M., Chishti, A. H., Crompton, A., Chan, A. C., Anderson, J. M., Cantley, L. C. (1987) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science.* 275, 73-77.
- Taliana, L., Benezra, M., Greenberg, R.S., Masur, S.K., Bernstein A.M. (2005) ZO-1: Lamellipodial localization in a corneal fibroblast wound model. *Invest. Ophthalmol. Vis. Sci.* 46, 96-103.
- Torrado, M., Senatorov, V. V., Trivedi, R., Fariss, R. N., Tomarev, S. I. (2004) Pdlim2, a novel PDZ-LIM domain protein, interacts with alpha-actinins and filamin A. *Invest Ophthalmol. Vis. Sci.* 45, 3955-3963.
- Van der Flier, A., Sonnenberg A. (2001) Structural and functional aspects of filamins. *Biochemica et Biophysica Acta.* 1538, 99-117.

- Wachsstock, D. H., Wilkins, J. A., Lin, S. (1987) Specific interaction of vinculin with alpha-actinin. *Biochem. Biophys. Res. Commun.* 146, 554-560.
- Wiche, G. (1998) Role of plectin in cytoskeleton organization and dynamics. *J. Cell Sci.* 111, 2477-2486.
- Wittchen, E. S., Haskins, J., Stevenson, B. R. (1999) Protein interactions at the tight junction: actin has multiple binding partners, and ZO-1 forms independent complexes with ZO-2 and ZO-3. *J. Biol. Chem.* 274, 35179-35185.
- Yamagata, N., Shyr, Y., Yanagisawa, K., Edgerton, M., Dang, T. P., Gonzalez, A., Nadaf, S., Larsen, P., Roberts, J. R., Nesbitt, J. C., Jensen, R., Levy, S., Moore, J. H., Minna, J. D., Carbone, D. P. (2003) A training-testing approach to the molecular classification of resected non-small cell lung cancer. *Clin. Cancer Res.* 9, 4695-4704.
- Yokoyama, S., Tachibana, K., Nakanishi, H., Yamamoto, Y., Irie, K., Mandai, K., Nagafuchi, A., Monden, M., Takai, Y. (2001) Alpha-catenin-independent recruitment of ZO-1 to nectin-based cell-cell adhesion sites through afadin. *Mol. Biol. Cell.* 12, 1595-1609.

Part 4:

Development of off-line HPLC MALDI-MS/MS methods for high-content phosphoproteomics using LC retention time prediction

4.0 Author's Contributions

The following work was performed after I was awarded a National Science and Engineering Research Council (NSERC, Canada) and National Science Council (NSC, Taiwan) Graduate Student Fellowships to conduct collaborative research in Taipei, Taiwan in 2006.

I initiated, designed and directed the execution of the majority of the work presented (outside of actual programming which was kindly provided by Mrs. H.-Y. Hsieh) towards the preparation of this manuscript. Dr. C.-C. Chou, provided MALDI-TOF-TOF training, helped to locate reagents and prepared and coordinated some of the standard tryptic mixtures and material used for retention time studies. C.-C. Chou and K.-H. Khoo provided helpful discussions and critical analysis. All research equipment and materials were provided by the National Core Facilities for Proteomics located at the Institute for Biological Chemistry, Academia Sinica (Taipei, Taiwan).

**Targeted identification of phosphorylated peptides by Off-line HPLC-
MALDI-MS/MS using LC Retention Time Prediction***

Vincent C. Chen¹, Chi-Chi Chou², Hsin-Yu Hsieh², Helene Perreault¹, Kay-Hooi Khoo²

¹University of Manitoba, Department of Chemistry, 144 Dysart Rd., Winnipeg, MB, Canada.

²National Core Facilities for Proteomics, Institute for Biological Chemistry, Academia Sinica, Taipei, Taiwan.

*In preparation

4.1 Abstract

Protein phosphorylation is a post-translational modification which plays an important role in cell regulation and signal transduction. Because of its biological relevance, a considerable amount of interest has been paid to the development of efficient techniques for phosphopeptide analysis. Although advances in MS control have enabled the high-throughput discovery of proteins from limited amounts of sample, automated selection of MS/MS precursor ions based on intensity alone can significantly lower the detection of low abundance phosphopeptides. Based on the observation that the introduction of a phosphate moiety does not dramatically change peptide retention time in reverse phase chromatography, phosphopeptide specific MS/MS fragmentation attempts based on LC retention time and m/z were evaluated using a standard protein mixture, *in vitro* phosphorylated myelin basic protein and immunoprecipitates of proteins from S2 cells. The development of these methods will not only permit the targeted identification of protein phosphorylation sites, but also allow the in-depth analysis of the dynamic events linked to the post-translational modification. Results indicated the vast majority (98%) of phosphopeptides identified eluted within a +/-4 minute window of predicted LC elution time. While studies presented here are primarily proof-of-concept in nature, data suggest that the use of LC retention time prediction could be a valuable constraint for the identification of phosphopeptides within a set of off-line LC deposited sample spots.

4.2 Introduction

Recent developments in mass spectrometry and maturation of proteomic technologies have provided researchers reliable and sensitive methods to identify proteins (Chen et al., 2006). In most proteomic shotgun studies, LC separated peptides are continuously infused into the emitting tip of an ESI-type ion source (e.g. ESI, nanospray) for MS/MS fragmentation and identification. Despite the development of MS/MS instrumentation with faster rates of acquisition (Wolters et al., 2001; Plumb et al., 2006; Luo et al., 2005), significant challenges related to the identification of low abundance proteins still exist. In a typical data-dependent scan, the peptides eluting from an LC column are first examined in MS mode and the most intense signals are automatically selected for MS/MS fragmentation. While ion intensity-based 'triggering' is suitable for the identification of analytes with high to moderate concentrations, low abundance molecules and peptides carrying post-translational modifications (PTMs) are often underrepresented within lists of identified proteins. While the introduction of automation has revolutionized proteomics, MS/MS acquisitions based on ion intensity alone can significantly lower detection limits of phosphopeptides even when the majority of analytes in question are well within instrumental dynamic range (Picotti et al., 2007).

With the recent development of MALDI instrumentation capable of MS/MS (Vestal et al., 2005; Ens and Standing, 2005), significant interest has been directed towards the off-line LC separation and deposition of peptides

(Chen et al. 2004; Krokhin et al. 2005; Snovida et al. 2006). Due to the nature of the ion source, the “off-line” coupling of LC to MALDI is commonly addressed by mixing LC effluent with a matrix and depositing discrete fractions onto a target surface. Although the addition of a fraction collection step can be viewed as an inconvenience, this method offers a number of distinct advantages including the option to completely analyze chromatographic runs in MS-mode prior to MS/MS acquisition. Further, as only a small portion of the static sample is used to generate a MALDI spectrum, a single LC run can be repeatedly analyzed in MS or MS/MS modes without the need of replicating sample injection(s). While the protocol is not a substitute for high-quality chromatographic separations, the off-line protocol enhances analyses by permitting the selection of precursor ions at maximum intensity - regardless of the presence or density of co-eluting peptides. Finally, as any parameter (including: m/z , abundance, retention time, etc.) alone or in combination can be used for precursor ion selection, the off-line protocol offers a flexible yet highly sensitive platform for the characterization of protein post-translational modifications (PTM).

Phosphorylation is a PTM that reversibly modulates inter- and intra-protein interaction domains/motifs which control protein-protein interactions and signal transduction events within the cell (Hunter, 2000; Pawson et al. 2005; Pawson 2007). Consequently, the experimental determination of phosphorylation status of a given protein has proven to be an important but

challenging task. While phosphoproteomics has been traditionally synonymous with proteomics, subtle but important differences exist. For example, proteomics is generally performed with the aim of identifying unknown proteins, phosphopeptide analyses are often conducted after a protein or a pool of protein candidates is known. In other words, while proteomic studies are largely performed in “hypothesis-generating” mode, researchers are often called upon to identify/track protein phosphorylation status in a “hypothesis-driven” manner (e.g. “protein X is hypothesized to be phosphorylated on all or a set of S/T/Y residues”). While the off-line LC-MS protocol does permit the search for peptide masses (e.g. peptide + phosphate/80Da), the persistence of multiple analytes having identical or near-identical masses with phosphopeptides in question, will still lead to higher rates of sample consumption. To address these challenges, more proficient methods for phosphopeptide analysis are required.

For a given set of separation conditions (e.g. solvent, gradient, ion pair, temperature etc.), LC retention time is directly dependent on the chemical features of the analyte (e.g. amino acid composition, sequence, hydrophobicity and structure) (Brown et al. 1982; Tripet et al. 2007). The LC retention data has received considerable interest from the proteomic research community because of its potential to help increase quality and quantity of protein identifications (Petritis et al. 2003; Petritis et al. 2005). The ability to predict LC elution time is a resource for predicting the location of a peptide

within a set of off-line HPLC-MALDI deposited spots (Krokhin 2006^a; Krokhin 2006^b, Krokhin et al. 2005). Based on the observation (Chen et al., 2004; Wu et al., 2006) that the introduction of a phosphate moiety does not dramatically change C18 LC retention time, a workflow incorporating the parameter was developed to increase the efficiency and sensitivity of phosphopeptide fragmentation attempts by off-line HPLC-MALDI. Proof-of-concept for phosphopeptide LC retention time prediction was demonstrated using a protein mixture containing bovine serum albumin, myoglobin, ovalbumin, fetuin, alcohol dehydrogenase and phosphorylated α - and β -caseins. To further extend the evaluation dataset and validate the workflow, the study also included the analysis of *in vitro* phosphorylated myelin basic protein. Finally, to replicate “real world” conditions, immunoprecipitates of S2 cell proteins were analyzed using the newly established protocol, resulting in the identification of several novel phosphorylation sites within Abelson-interacting protein, ABI.

4.3 Experimental

4.3.1 Materials.

All reagents including bovine serum albumin (BSA), bovine α - and β -casein, myelin basic protein (MBP), fetuin, horse myoglobin, chicken ovalbumin, yeast alcohol dehydrogenase (ADH), monoclonal anti-phosphotyrosine-agarose, trifluoroacetic acid (TFA), trichloroacetic acid (TCA), formic acid, α -cyano-4-hydroxycinnamic acid (CHCA) matrix, dithioerythritol (DTT), iodoacetamide and ammonium bicarbonate were obtained from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise stated. Protein kinase A was obtained from New England BioLabs Inc. (Ipswich, MA, USA). Sequence grade trypsin (Promega, Madison, WI) was used for protein digestions and all LC mobile phases were prepared with deionized (18 M ohm/cm) water using HPLC grade acetonitrile (ACN, Merck, Darmstadt, Germany).

4.3.2 Off-line high performance liquid chromatography and MALDI fraction collection.

Off-line high performance liquid chromatography and MALDI fraction collection were performed using a nano-flow Ultimate HPLC with a UV micro flow detector (LC-Packings/Dionex, Ultimate™) and a Probot MALDI deposition/fraction collection system (LC-Packings/Dionex, San Francisco, Ca). Separations were performed on 22 cm (L), 75 μ m (ID) micro columns containing 100 Å, 5 μ m Nucleosil C18 (Sorbent Technologies, Atlanta, Georgia) reverse phase material. All samples were injected using a 5 μ L

sample loop with the HPLC's manual injection port. Separation quality was continuously monitored by UV absorption at 254 nm using a nanovolume flow cell (LC-Packings, 3 nL, 20 μ m, Figure 4.1). Fused silica capillary tubing (OD 365 μ m, ID 75 μ m) obtained from LC-Packings (Amsterdam, The Netherlands) was used for all HPLC fluid transfer lines. Matrix solution, comprised of 10 mg/mL α -cyano-4-hydroxycinnamic acid in 80% acetonitrile/0.1% TFA, was prepared and continuously infused in-line using a peek micro-T junction (LC-Packings) at a rate of 0.7 - 0.8 μ L/min prior to sample deposition. Fractions were collected on an Applied Biosystems 198 spot (24 x 24) MALDI surface at 1 minute intervals. Two mobile phases, solvent A: 95% water, 5% ACN (0.1% TFA) and solvent B: 95% ACN, 5% H₂O (0.1% TFA) were used to generate separation gradients. Prepared mobile phases were sonicated under vacuum for 10 minutes and purged with helium gas for 15 minutes prior to and during liquid chromatography. Two different linear HPLC gradient conditions, with 200 nL/minute flow rates were used: Condition 1 was a linear gradient of 0-70% solvent B over 90 minutes; and Condition 2 was a gradient of 0-100% solvent B over 80 minutes.

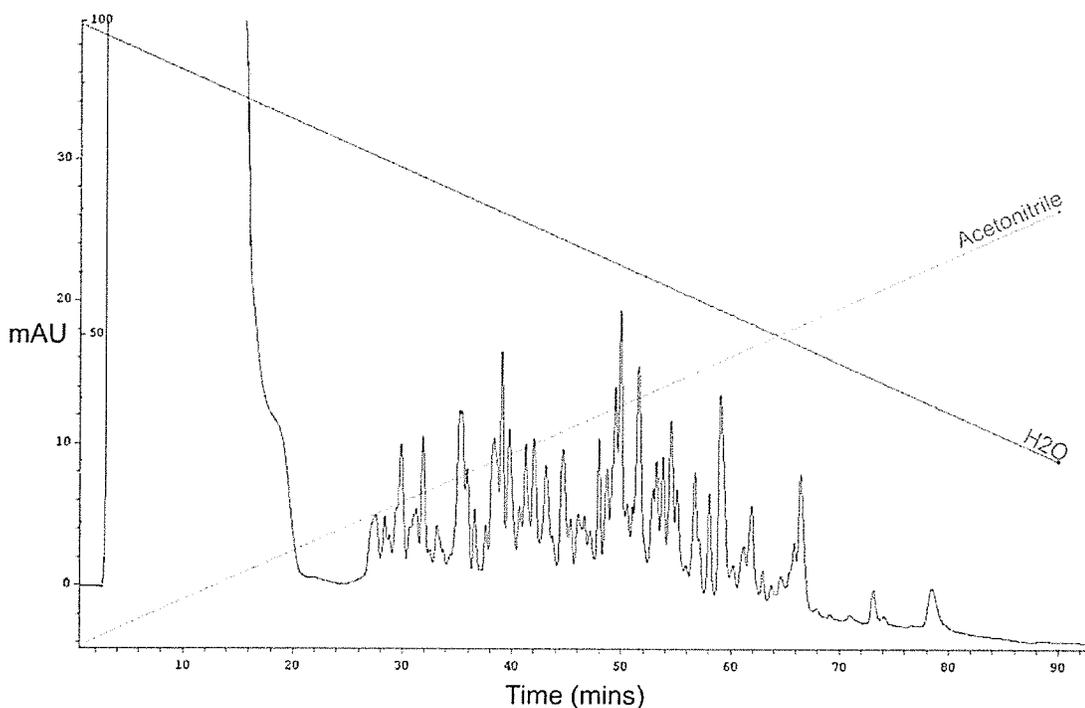


Figure 4.1: HPLC-UV trace. UV monitoring (254 nm) of the reverse phase separation of the 7-protein digest mixture using Condition 1.

4.3.3 Mass spectrometry

All MS and MS/MS spectra were acquired on a MALDI-TOF/TOF 4700 Proteomics Analyzer™ mass spectrometer (Applied Biosystems, Framingham, MA) equipped with an Nd:YAG laser (355 nm wavelength, 500-ps pulse, and 200 Hz repetition rate in both MS and MS/MS modes) with 1000 and 5000 shots accumulated in positive ion mode in MS and MS/MS modes respectively.

In MS operation, all fraction spectra were acquired for the mass range of 800–4000 Da and automatically processed by the Applied Biosystems

4700 Explorer™ software which produced individual PMF peak list files. In MS/MS operation, 1 keV collisions with argon were used to generate the high-energy CID spectra using a source voltage of 8 kV, a collision cell voltage of 7 kV, and a second accelerating voltage of 15 kV. GPS Explorer™ v2.0 software (Applied Biosystems) integrated to a 4 processor Mascot (v2.1, Matrix Science, UK) server. Spectra of phosphorylated and non-phosphorylated peptides were manually validated and processed with Mascot Distiller™ (Matrix Science) or Data Explorer™ (Applied Biosystems) programs and automated protein identifications were carried out using the NCBI nr database (www.ncbi.nlm.nih.gov/entrez).

4.3.4 Workflow and Data-dependent analysis.

HPLC fractions were analyzed in single MS mode and m/z with signal-to-noise (S/N) values greater than 10 were assigned and compiled as batch peptide-mass fingerprints (PMF) files by the instrument software. Tryptic peptides of myoglobin were used as internal standards to calibrate the HPLC runs for retention time prediction as previously described (Krokhin, 2006^a). Calibration peptides eluting over multiple fractions were given a single weighted-average fraction number based on relative peak area ratios (Table 4.2). The *SSRCalc* (Krokhin, 2006^a, www.systemsbio.ca) was used to determine the relative hydrophobicities of all peptides including the calibration peptides. Calibration values were plotted against observed weighted fraction number (time in minutes) using functions within Microsoft Excel™ (Redmond,

Table 5.1: Calibration and LC retention time data for phosphorylated protein α - and β -casein within the 7 protein mixture.

m/z (predicted)	m/z (Exp)	Fraction # (predicted) $y = 0.7582x + 33.913$	Fraction # (Exp)	Sequence – Protein	Rel. Hydropho.	PTM
1466.61	1466.587	52	51/52	TVDMESEVFTK- α Casein	23.84	1PO4
1594.71	1594.685	51	49/50	TVDMESEVFTKK- α Casein	22.13	1PO4
1660.795	1660.751	52	51/52/53/54	VPQLEIVPNSAEER- α Casein	23.60	1PO4
1832.83	1832.756	57	58/59	YLGEYLIVPNSAEER- α Casein	30.76	1PO4
1847.69	1847.710	50	50	DIGSESTEDQAMEDIK- α Casein	21.37	1PO4
1951.953	1951.887	54	53/54/55	YKVPQLEIVPNSAEER- α Casein	25.82	1PO4
2061.829	2061.757	46	44/45/46/47	FQSEEQQTTEDELQDK- β Casein	15.44	1PO4
1086.562	1086.493	41	40/41/42	HLKTEAEMK-Myoglobin	13.4	
1271.664	1271.648	50	49/50/51	LFTGHPETLEK-Myoglobin	23.1	
1360.759	1360.705	54	55	ALELFRNDIAAK- Myoglobin	29.98	
1378.842	1378.775	Calibration	61/62/63 (61.4)*	HGTVVLTAALGGILK- Myoglobin	40.29	
1502.670	1502.621	Calibration	45/46/47 (46.0)*	HPGDFGADAQGAMTK- Myoglobin	20.18	
1506.937	1506.867	Calibration	59/60 (60.0)*	KHGTVVLTAALGGILK- Myoglobin	38.79	
1606.855	1606.832	Calibration	50/51/52 (51.2)*	VEADIAGHGQEVLR- Myoglobin	27.29	
1661.854	1661.762	Calibration	53	LFTGHPETLEKFDK-Myoglobin	29.06	
1853.962	1853.856	Calibration	43/44/45 (43.5)*	GHHEAELKPLAQSHATK- Myoglobin	18.04	
1982.057	1981.941	41	43	KGHHEAELKPLAQSHATK-Myoglobin	16.14	

*Peptides used for calibration with the weighted values used in brackets. **Bold** indicates actual fractions used for MS/MS. Myoglobin peptides not used for calibration were used as positive controls.

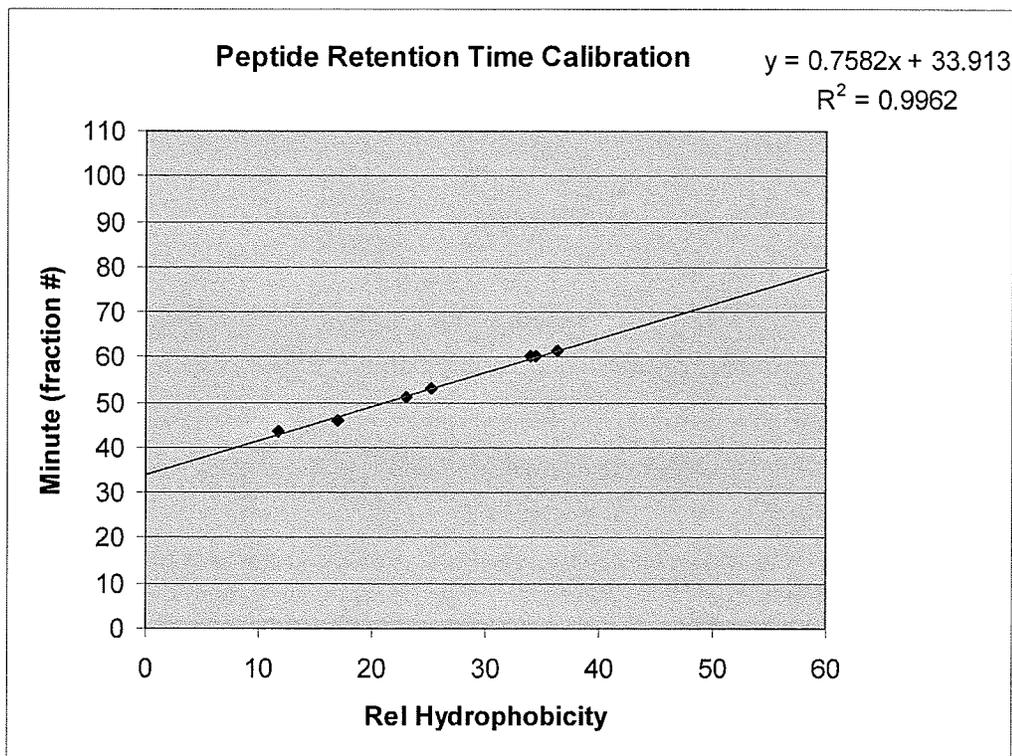


Figure 4.2: Retention time calibration using myoglobin peptides. All sequences of myoglobin were confirmed by MS/MS. R^2 values ranging between 0.95-0.99 were observed over the course of the study.

WA) to determine best-fit line, slope, y-intercept and R^2 values (Table 4.1, Figure 4.2). The slope and y-intercept values, along with peptide sequences of interest, were then used by the *SSRCalc* to calculate LC elution time by linear regression. Peptide sequences were generated using the following conditions and parameters. Theoretical digests were generated using MS-Digest (<http://prospector.ucsf.edu/>) selecting the following parameters: trypsin as the digest enzyme, a maximum of 2 missed cleavages, iodoacetamide modified cysteine (carboxyamidomethyl cysteine), m/z 700 - 5000 with the

partial phosphorylation of serine (S), threonine (T), and tyrosine (Y) residues. Sequences, PTM and m/z values determined by MS-Digest were exported to into an Excel™ data worksheet, with columns corresponding to theoretical values for native amino acid sequence, peptide modification(s) and m/z . As *SSRCalc* (version 3.1) does not recognize modified amino acids, native sequences carrying theoretical PTMs were noted within the spread sheet prior to *SSRCalc* web submission and retention time prediction. Predicted retention times were then rounded to the closest whole integer and designated a theoretical fraction number. Theoretical sequences, modifications, protein of origin, m/z , and *SSRCalc* values were tabulated in Excel™ and referenced for MS/MS (inclusion lists) for MALDI-MS/MS analysis.

To help identify fractions containing phosphorylated and non-phosphorylated peptide candidates, an in-house executable file called *SearchPeak* (Figure 4.3) was written in Microsoft Visual Basic™. The program operates by matching data in PMF batch files with theoretical values for peptide retention time within the peptide retention time worksheet. The program allowed for the rapid assessment of MS data prior to MS/MS acquisition within user-defined tolerances for mass and predicted *SSRCalc* fraction number. For the purpose of this study, retention time prediction was evaluated using tolerances set to ± 0.5 Da and ± 4 fractions which was determined adequate during preliminary studies (data not shown).

Search Peak

Search Peak List

C:\Documents and Settings\V. Chen\Desktop\20060801-LC.ms-7prtens_TFA PeakList Folder

C:\Documents and Settings\V. Chen\Desktop\7protein_trfalist.txt PeakFile (.txt)

MW Range (+/-) 0.2 Spot Range (+/-) 2 Search Clear

UnModified Match Select All Copy

Modified Match Select All Copy

958.4842_958.4007_48_043_EALDFFAR_Akchoh
 958.4842_958.37823_48_049_EALDFFAR_Akchoh
 1013.5995_1013.5614_43_043_AHELLINVK_Akchoh
 1251.6697_1251.6224_42_043_SISVGSYVGNR
 1251.6697_1251.6124_42_044_SISVGSYVGNR
 1433.7892_1433.665_47_046_YVGLSSLPEYEK
 2019.0697_2018.9449_46_046_LPLVGGHEGAG
 2019.0697_2018.917_46_047_LPLVGGHEGAGY
 2477.1461_2476.9976_47_046_YSGVCHTDLHA
 2477.1461_2476.9661_47_047_YSGVCHTDLHA
 2477.1461_2476.9478_47_048_YSGVCHTDLHA

2061.829_2061.8511_36_036_FQSEEQDQTEC
 2061.829_2061.8704_36_037_FQSEEQDQTEC
 2061.829_2061.8326_36_038_FQSEEQDQTEC
 1660.7947_1660.7462_42_042_VPQLEIVPNSA
 1660.7947_1660.7327_42_043_VPQLEIVPNSA
 1951.953_1951.863_44_044_YKVPQLEIVPNSA
 1951.953_1951.8513_44_045_YKVPQLEIVPNSA
 1951.953_1951.8331_44_046_YKVPQLEIVPNSA

Output

“Non-modified” “PTM-modified”

+/- Da (M/z)

Tolerance Retention Time +/- (min)

Figure 4.3: *Search Peak* interface. Software used to identify phosphopeptides within user defined tolerances for mass and retention time. Software input included theoretical peptide sequence, m/z , predicted retention time within the “Peaklist folder”, and experimental m/z and fraction number within the “Peakfile (.txt)”. MS/MS candidates of peptides (Unmodified Match) and phosphopeptides (“Modified Match”) were predicted by cross-referencing values.

To account for the likelihood of phosphopeptides/peptides residing outside predicted values, all fractions were then searched '*in-silico*' using *Search Peak* (i.e. inclusion list based on *m/z* only) followed by MS/MS validation. All confirmed phosphopeptides (*m/z*, sequences) and corresponding fraction numbers were compiled for further evaluation for the purpose of this study.

4.3.5 Seven-protein tryptic digest mixture.

Bovine serum albumin (BSA), myoglobin, ovalbumin, fetuin, alcohol dehydrogenase and phosphorylated α - and β -caseins were individually prepared as 10 mg/mL solutions in 25 mM ammonium bicarbonate. Volumes corresponding to equal molar amounts were aliquoted, mixed, reduced with DTT (65 mM, 1h, at 37°C) and alkylated using iodoacetamide (100 mM, 1h in the dark at room temperature). Proteins were digested in solution using sequence-grade modified trypsin in ammonium bicarbonate (25 mM, 16 h, at 37 °C). Resulting peptides were then lyophilized to dryness and resuspended in 0.1% TFA to a final concentration of 0.2 pmole/ μ L. A standard volume of 5 μ L was injected for all off-line HPLC-MALDI-MS runs. Analyses were performed in triplicate (Table 4.2) and MS/MS spectra were manually interpreted by visual inspection (Figure 4.4).

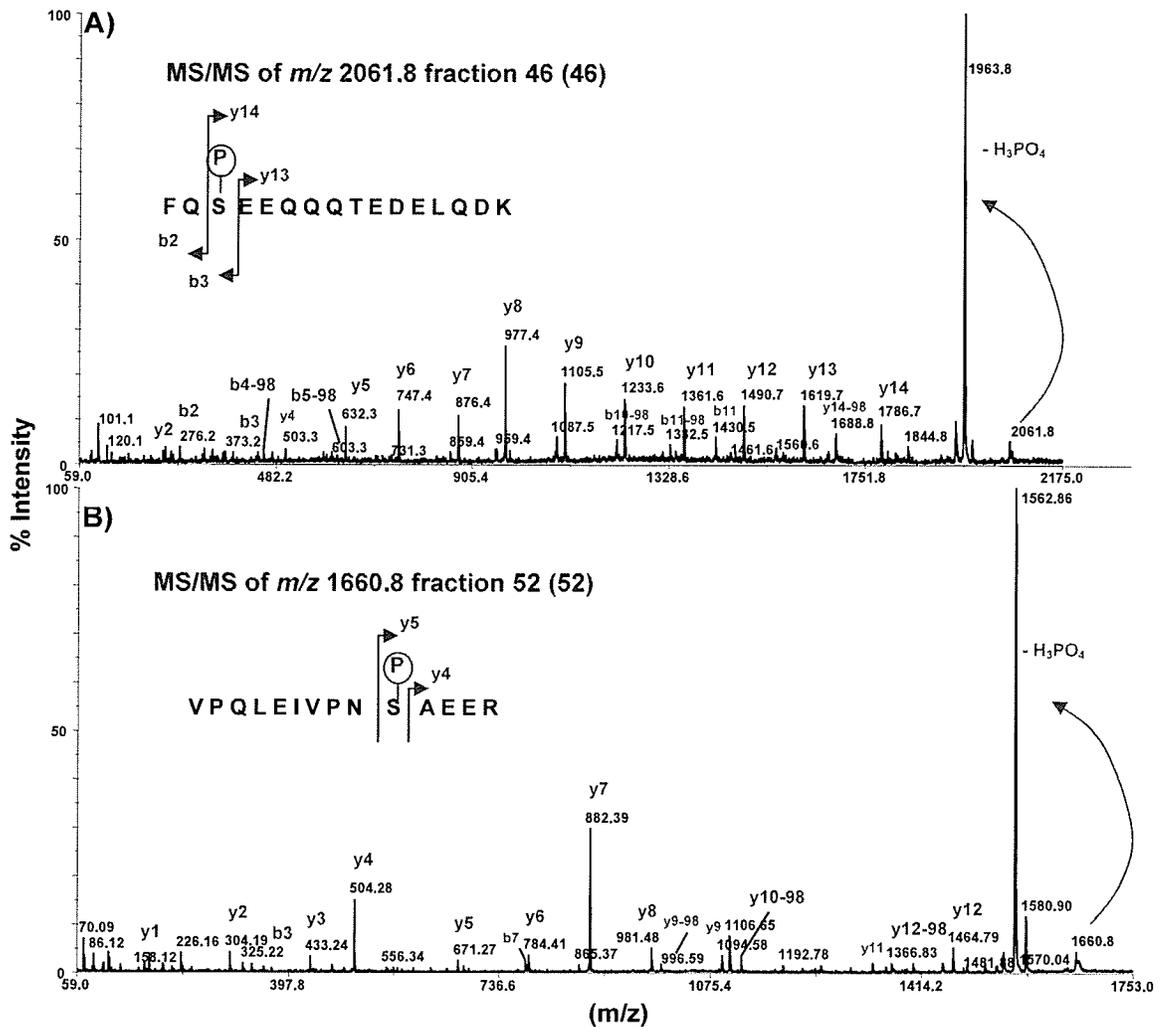


Figure 4.4: MS/MS of casein phosphopeptides from (A) β -casein (FQpSEEQQQTEDELQDK) and (B) α -casein (VPQLEIVPNpSAEER) within the 7-protein digest mixture (HPLC-MALDI-MS run 1). Experimentally observed fraction and theoretical retention time fraction numbers (in brackets) are indicated. Diagnostic losses of phosphoric acid from the parent masses are labeled (green).

Table 4.2: LC retention data of α - and β -casein phosphopeptides observed.

Phosphopeptide sequence	Number of phosphoryl groups	[M+H] ⁺	Run 1 Δ fraction (minute)	Run 2 Δ fraction (minute)	Avg. Δ fraction (minute)
¹⁵³ TVDME _p STEVFTK ¹⁶⁴	1	1466.6	0	-1	-0.5
¹⁵³ TVDME _p STEVFTKK ¹⁶⁵	1	1594.7	-2	-3	-2.5
¹²¹ VPQLEIVPN _p SAEER ¹³⁴	1	1660.3	0	0	0
¹⁰⁴ YLGEYLIVPN _p SAEER ¹¹⁹	1	1832.8	+1	+1	+1
⁵⁸ DIG _p SE _p STEDQAMEDIK ⁷³	2	1927.7	0	+1	+0.5
¹¹⁹ YKVPQLEIVPN _p SAEER ¹³⁴	1	1951.4	0	0	0
³ FQ _p SEEQQQTEDELQDK ⁴⁸	1	2061.8	0	0	0

4.3.3 Myelin Basic Protein (MPB)

In vitro phosphorylated MBP was prepared by treating the protein with protein kinase A in 50 mM Tris-HCl buffer (pH 7.4) containing 10 mM MgCl₂ and 0.2 M ATP for 4h at room temperature and precipitated using TCA. Solutions of proteins and trypsin were prepared in aqueous ammonium bicarbonate buffer (25 mM, pH 8.5). Stock solutions of proteins (approximately 10 μ g), were reduced with DTT at 37 °C for 1hr, alkylated with iodoacetamide at 37°C for 1hr and treated overnight with trypsin at an enzyme-to-substrate ratio of 1:50 at 37°C as previously described in section 4.3.2. The digested products were then diluted with 0.1% TFA to a concentration of 2 pmole/ μ L and 5 μ L per injection was performed. LC retention time prediction was conducted as previously described. Phosphopeptide candidates falling within predetermined tolerances for *m/z*

and LC retention time were identified. Phosphorylated MBP peptides were identified, including phosphopeptides HRDpTGILDSLGR and NIVTPRpTPPPSQGK presented (Figure 4.5). In total ten phosphopeptides of MBP were characterized over three replicates (Table 4.3).

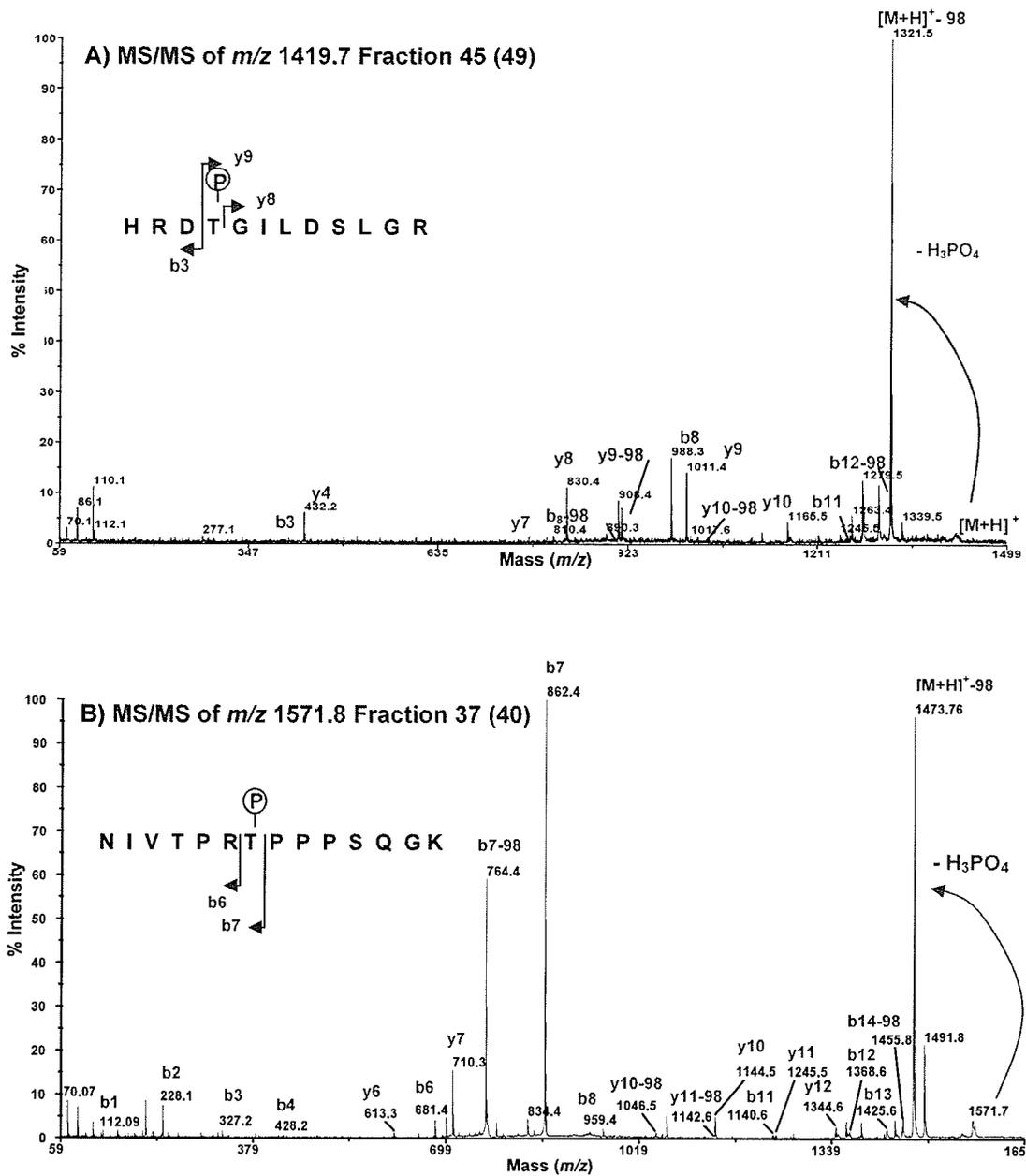


Figure 4.5: MALDI-MS/MS spectra of tryptic phosphopeptides of myelin basic protein m/z (A) 1419.7 (HRDpTGILDSLGR) and (B) 1571.8 (NIVTPRpTPPPSQGK). Experimentally observed fraction number and theoretical retention time fraction numbers (in brackets) are indicated. Diagnostic losses of phosphoric acid from parent masses are labeled (green).

Table 4.3: LC retention time variance for tryptic phosphopeptides from myelin basic protein.

Sequence	number of phosphoryl groups	[M+H] ⁺ (theoretical)	run 1 deviation (minute)	run 2 deviation (minute)	run 3 deviation (minute)	mean bias SD (inter-run)	Accuracy mean relative error (%)
phosphopeptides							
¹⁵⁵ LGGRD <p>SR¹⁶¹</p>	1	840.4	+2	+1	-1	1.5	1.9
¹⁰⁵ GRGL <p>SLSR¹¹²</p>	1	925.5	-2	-1	-2	0.6	-2.7
or ¹⁰⁵ GRGLSLS <p>SR¹¹²</p>							
⁶⁴ T <p>THYGSLPQK⁷³</p>	1	1211.6	0	-1	-5	2.6	-4.3
⁵² RG <p>SGKDGHHAAR⁶³</p>	1	1328.6	1	0	0	0.6	1.0
³⁰ HRD <p>TGILD SLGR⁴¹</p>	1	1419.7	1	-1	-4	2.5	-2.6
¹⁵⁵ LGGRD <p>SRSGSPMAR¹⁶⁸</p>	1	1526.7	1	-4	0	2.6	-1.4
or ¹⁵⁵ LGGRDSR <p>SGSPMAR¹⁶⁸</p>							
or ¹⁵⁵ LGGRDSRSG <p>SPMAR¹⁶⁸</p>							
⁹¹ NIVTPR <p>TPPPSQK¹⁰⁴</p>	1	1571.8	0	0	-4	2.3	-3.0
¹¹³ F <p>SWGAEGQKPGFGYGGR¹²⁹</p>	1	1880.8	2	1	-3	2.6	0.2
¹⁰⁷ GLSLSRF <p>SWGAEGQKPGFGYGGR¹²⁹</p>	1	2494.2	0	0	-3	1.7	-1.7
¹⁰⁷ GL <p>SLSRF SWGAEGQKPGFGYGGR¹²⁹</p>	2	2574.1	1	0	-2	1.5	-0.5

Accuracy = [(exp. retention time – theoretical retention time)/ theoretical retention time] x100%. Accuracy presented as a mean relative error (%) was determined by averaging over each of the three runs.

4.3.7 Anti-phosphotyrosine immunoprecipitation.

Drosophila S2 cell cultures were a kind gift from Dr. T. Meng (Institute for Biological Chemistry, Academia Sinica). Homogenates of S2 cells were prepared in lysis buffer (20 mM Tris, pH 7.4, 15 mM NaCl, 1% Triton X-100, 1 mM EDTA, 1 mM sodium orthovanadate) in the presence of protease inhibitors. Cleared supernatants were incubated overnight at 4°C with 40 µL of monoclonal anti-phosphotyrosine-agarose antibody. Immunoprecipitates were washed 3 times in ice cold lysis buffer minus protease inhibitor. Phosphoproteins were then released by incubating the beads with 0.2 M phenylphosphate (in lysis buffer) for 1 h. Proteins were then reduced and alkylated as previously described. Enzymatic digestion was then carried out using sequence-grade modified trypsin at an enzyme-to-substrate ratio of 1:50 at 37 °C. Phosphoproteins were lyophilized to dryness and resuspended in 10 µL of 0.1% TFA. IP material was split into two portions (5 µL each), and the first sample was used to establish protein candidates by MS/MS automation (using the “top 5 triggering” function; i.e. the five most abundant intensities per fraction were subjected to MS/MS). The automated acquisition led to the identification of over 60 proteins (found in supplementary information, Section 7.0) including SCAR, heterogeneous nuclear ribonucleoprotein, SH3PX1 and Abelson-interacting protein (ABI) (Table 4.4).

Table 4.4: Top 5 proteins identified within anti-pY immunoprecipitates of S2 cell proteins by automated (intensity-driven) data-dependent analysis.

1. gij11042 hrp48.1 [Drosophila melanogaster]				
Observed	Mr(expt)	Delta	Score	Peptide
645.23	644.23	-0.05	25	YFCR+Carbamidomethyl Cys
684.31	683.31	-0.03	22	TFFNR
1032.47	1031.46	-0.06	35	TFFNRYGK
1197.53	1196.52	-0.07	58	TIDPKPCNPR + Carbamidomethyl Cys
1197.57	1196.56	-0.04	14	TIDPKPCNPR + Carbamidomethyl Cys
1206.49	1205.48	-0.06	130	QEGASNYGAGPR
1206.52	1205.51	-0.03	13	QEGASNYGAGPR
1206.53	1205.53	-0.02	23	QEGASNYGAGPR
1453.63	1452.63	-0.09	61	VTEVVIMYDQEK
1717.83	1716.82	-0.08	107	VFLGGLPSNVTETDLR
1769.69	1768.68	-0.08	170	DGSGGQNSNNSTVGGAYGK
1898.82	1897.82	-0.09	166	AQAWATGGPSTTGPVGGMPR
1898.87	1897.86	-0.05	9	AQAWATGGPSTTGPVGGMPR
1898.89	1897.88	-0.03	51	AQAWATGGPSTTGPVGGMPR
1936.89	1935.89	-0.08	126	LFVGGLSWETTQENLSR
1953.83	1952.82	-0.09	72	FGDIIDCVVMKNNESGR + Carbamidomethyl Cys
2122.02	2121.01	-0.07	50	GKLFVGGLSWETTQENLSR
2299.93	2298.93	-0.11	92	GFGFLSFEEESSVEHVTNER
2608.92	2607.91	-0.11	146	SGSEYDYGGYGSYDYDYSNYVK
2910.34	2909.33	-0.09	109	GFGFVTFADPTNVNHVLQNGPHTLDGR
2910.36	2909.35	-0.07	56	GFGFVTFADPTNVNHVLQNGPHTLDGR
3153.42	3152.41	-0.14	16	SRGFGFVTFADPTNVNHVLQNGPHTLDGR
3796.4	3795.39	-0.16	262	SGSEYDYGGYGSYDYDYSNYVKQEGASNYGAGPR

2. gij19921124 SCAR CG4636-PA [Drosophila melanogaster]				
Observed	Mr(expt)	Delta	Score	Peptide
992.5	991.5	-0.07	18	IRVPHNTR
1014.56	1013.56	-0.04	26	AIRDGITLR
1109.48	1108.47	-0.07	23	MLPPFHDPR
1139.53	1138.53	-0.05	66	VFDQQIFSR
1139.54	1138.53	-0.05	59	VFDQQIFSR
1139.55	1138.54	-0.03	13	VFDQQIFSR
1139.56	1138.55	-0.02	18	VFDQQIFSR
1237.58	1236.57	-0.07	49	KMLPPFHDPR
1243.54	1242.53	-0.07	101	HAEDVFGELAR
1310.69	1309.68	-0.05	107	NAAPLDVASILAR
1797.75	1796.74	-0.08	25	FYTDPNYFFELWR
1826.86	1825.86	-0.1	112	ALVHGETLMPNNVIYR
2015.95	2014.94	-0.1	154	VTQLDSTVEEVPLTDITR
2400.09	2399.08	-0.12	87	SLANGEMQQPGQQNGVPHIVAPK
2441.24	2440.24	-0.11	45	LAIKVTQLDSTVEEVPLTDITR
2994.45	2993.44	-0.08	175	SVYQQDELQSVELETVTNTTLTNIIR
3110.41	3109.4	-0.13	7	NKPRPSQPPPAPPSNGSGGTPTASNANTPTR
3869.77	3868.77	-0.18	80	MSPPNAAAPPPPPPPVEEGMGSGNQHTLRPHQILPK

Table 4.4 continued.

3. gi 24650831 Heterogeneous nuclear ribonucleoprotein at 98DE CG9983-PA, isoform A [Drosophila melanogaster]				
Observed	Mr(expt)	Delta	Score	Peptide
631.27	630.26	-0.05	27	AHFEK
949.37	948.37	-0.05	53	GGGNFNNNR
1053.53	1052.52	-0.05	86	LFIGGLDYR
1065.52	1064.51	-0.07	38	SRPHKIDGR
1073.42	1072.42	-0.05	80	QNDQQGGGGGR
1169.46	1168.45	-0.08	71	MQPYQGGGGFK
1181.62	1180.61	-0.05	79	KLFIGGLDYR
1482.67	1481.66	-0.07	101	ALPKQNDQQGGGGGR
1554.62	1553.61	-0.09	93	QNDQQGGGGGRGGPGGR
1554.65	1553.64	-0.07	118	QNDQQGGGGGRGGPGGR
1627.8	1626.79	-0.07	44	WGNIVDVVMKDPR
2017.85	2016.85	-0.08	37	GFGFITYSHSSMIDEAQK
2237.83	2236.82	-0.11	170	AGGGNQGNYGGNNQGFNNGGNNR
2331.05	2330.05	-0.1	69	GFAFVEFDDYDPVDKVV LQK
2851.03	2850.02	-0.12	53	GNMGGGNYGNQNGGGNWNNGGNNWGNNR
3306.22	3305.22	-0.15	14	AGGNRGNMGGGNYGNQNGGGNWNNGGNNWGNNR

4. gi 21355219 SH3PX1 CG6757-PA [Drosophila melanogaster]				
Observed	Mr(expt)	Delta	Score	Peptide
1249.5	1248.49	-0.05	78	SDVGEGWWEQK
1590.64	1589.64	-0.09	116	SDVGEGWWEQKNAR
1634.71	1633.71	-0.09	86	GVLNCFDPDIFSTHK + Carbamidomethyl Cys
1717.86	1716.85	-0.07	47	IGGIFIGIQAFGDQPK
2132.01	2131	-0.09	93	GVLNCFDPDIFSTHKGAIQK + Carbamidomethyl Cys
2180.03	2179.03	-0.07	129	GQIGLFPAAVVEVMSAAEAQK
2215.03	2214.03	-0.08	26	VDAQVELGTQFIHSM DVAVR
2291.11	2290.11	-0.09	106	LSASGATSVQVPDPFASPLPR
2703.17	2702.16	-0.12	154	AMYDFTGEPGSSELSIATGDVLSVTR
2703.19	2702.18	-0.09	33	AMYDFTGEPGSSELSIATGDVLSVTR

5. gi 17137386 Abelson Interacting Protein CG9749-PA [Drosophila melanogaster]				
Observed	Mr(expt)	Delta	Score	Peptide
1481.68	1480.67	-0.06	36	QSLRDSYTNLER
1736.96	1735.95	-0.09	93	IVAPINPEKPIKYVR
2444.18	2443.17	-0.11	67	TSTGSQLAPIVPEDQNLPGWVWK
2695.21	2694.21	-0.11	124	KPIDYSMLDEIGHGINS AQHSQVR
2695.25	2694.25	-0.07	170	KPIDYSMLDEIGHGINS AQHSQVR
2842.28	2841.27	-0.14	77	TPPVVNPPQVPSHYAPNYPIGHPKR + Phospho (Y)*
3075.54	3074.54	-0.07	52	TSTGSQLAPIVPEDQNLPGWVWKNFIEK
3121.46	3120.46	-0.15	40	GSSHGSVQSLLPPSVGPPPTTKPPTPPQMSR

*Phosphopeptide identified (red).

Of the MS/MS triggered analyses, only one phosphopeptide (ABI sequence: TPPVNPQVPSHYAPN_pYPIGHPKR, Mascot Score 77) was confidently identified by manual validation (Figure 4.6). Because it demonstrated phosphorylation, the ABI protein was the subject for further interrogation using the protocol previously described. However, as the introduction of calibration peptides would likely cause the lowered detection (i.e. ion suppression) of low abundance peptides, myoglobin peptides were not added. Instead, endogenous peptides of hrp48.1 and SCAR (originally identified during the initial proteomic survey, refer Table 4.4) were used for this purpose (Table 4.5 and Figure 4.7).

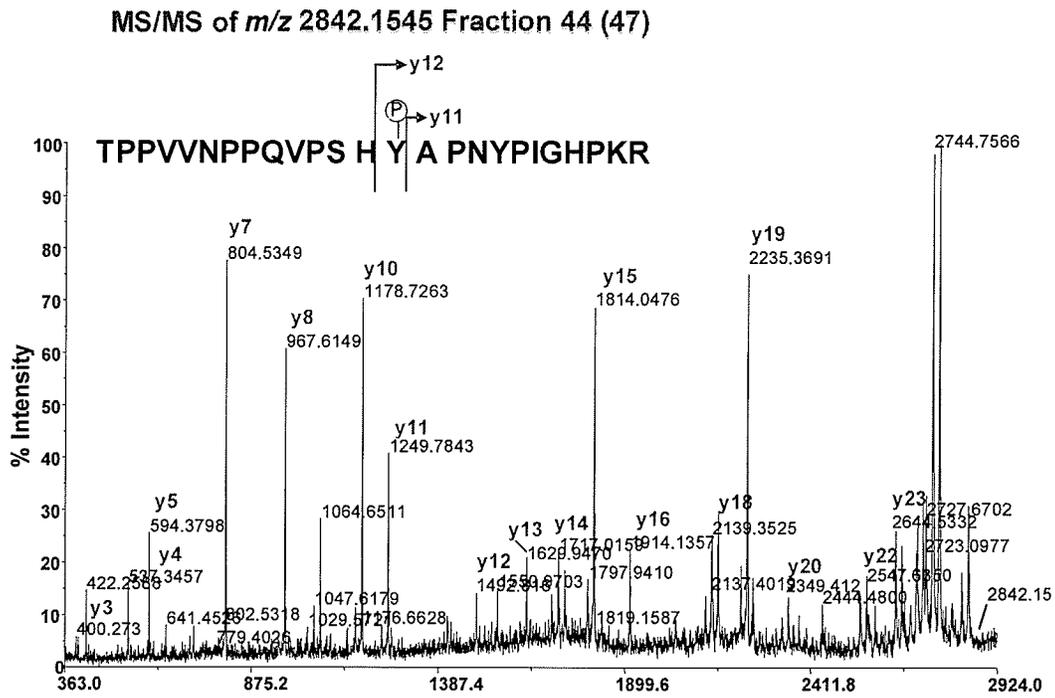


Figure 4.6: ABI phosphopeptide TPPVNPQQVPSH_pYAPNYPIGHPKR identified during the initial proteomic survey using automated data-dependent MALDI-MS/MS acquisitions. Experimentally observed fraction and theoretical fraction number (brackets) are indicated.

Table 4.5: SSR*Cal* calibration for Abelson-Interacting Protein

Protein	M/z	Relative Hydrophobicity	Fraction (exp)
Hrp48.1			
QEGASNYGAGPR	1206.55	5.39	34.15
TFFNRYGK	1032.527	21.08	43.4
DGSGGQNSNNSTVGGAYGK	1769.769	7.87	33.1
SCAR			
ALVHGETLMPNNVIYR	1826.959	29.14	47.3
SVYQQDELQSVELETVTNTTLTNIIR	2994.527	36.99	59

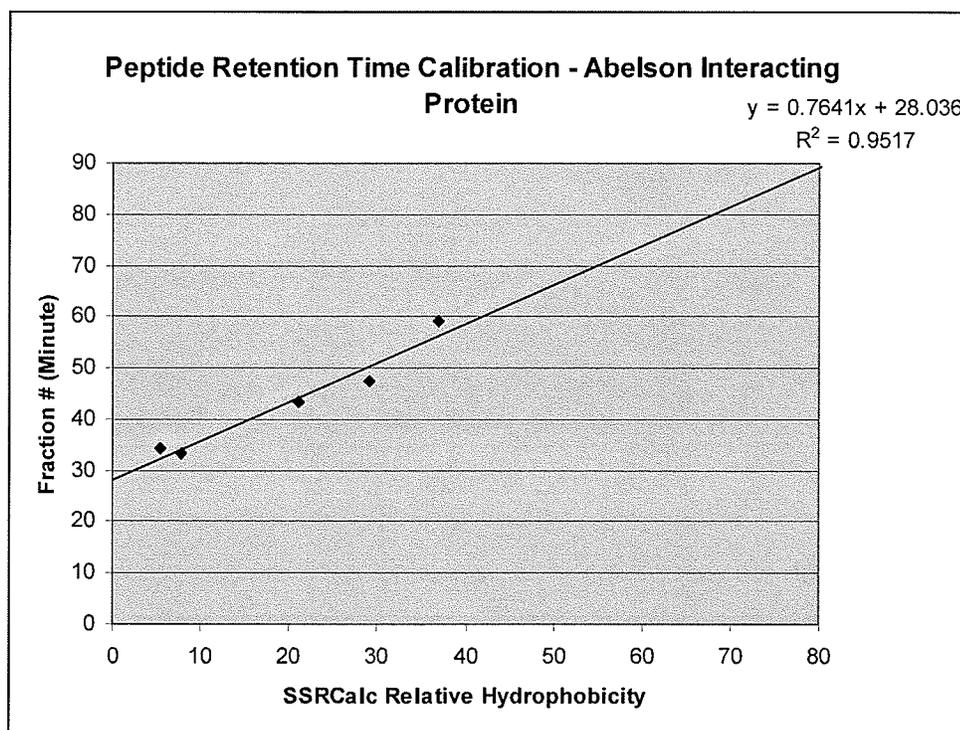


Figure 4.7: LC retention time calibration for peptides obtained from S2 Cell protein immunoprecipitates. This calibration ($R^2 = 0.95$) was subsequently used to predict the location of ABI phosphopeptides within the MALDI sample set.

Table 4.6: MS/MS inclusion list indicating fraction location, and *m/z* of candidate sequences of ABI.

<i>m/z</i>	<i>m/z</i> (exp)	Fraction #	Fraction # (exp.)	Abelson-interacting Protein (ABI) sequence	PTM
944.542	944.36	41	44	EIGVLTANK	
997.459	997.3556	38	37	DSYTNLER	
1077.43	1077.415	38	35	DSYTNLER	1PO4
1077.43	1077.417	38	36	DSYTNLER	1PO4
1293.68	1293.523	34	37	SVSNTGTLGKSSR	
1293.68	1293.514	34	38	SVSNTGTLGKSSR	
1453.74	1453.555	43	44	EIGVLTANKVSSR	1PO4
1453.74	1453.557	43	45	EIGVLTANKVSSR	1PO4
1454.59	1454.473	41	40	VADYCEDTYR	
1561.7	1561.533	42	42	QSLRDSYTNLER	1PO4
1737.04	1736.843	47	44	IVAPINPEKPIKYVR	
1737.04	1736.848	47	45	IVAPINPEKPIKYVR	
1871.93	1871.67	56	53	DDELSFQESSVLYVLK	
1871.93	1871.67	56	54	DDELSFQESSVLYVLK	
1882.79	1882.716	39	40	VADYCEDTYRADNK	
2089.91	2089.708	38	36	MSTASSTMTTTTTGGGAAGNER	
2089.91	2089.712	38	37	MSTASSTMTTTTTGGGAAGNER	
2089.91	2089.695	38	42	MSTASSTMTTTTTGGGAAGNER	
2365.15	2364.97	63	63	MLTETPMASENIMDELASLIR	
2365.15	2364.907	63	65	MLTETPMASENIMDELASLIR	
2365.15	2364.921	63	66	MLTETPMASENIMDELASLIR	
2365.15	2364.908	63	67	MLTETPMASENIMDELASLIR	
2433.03	2433.071	50	52	DSYTNLERVADYCEDTYR	
2444.28	2444.057	54	51	TSTGSQLAPIVPEDQNLPGWVWP	
2445.11	2445.012	63	62	MLTETPMASENIMDELASLIR	1PO4
2445.11	2445.032	63	63	MLTETPMASENIMDELASLIR	1PO4
2686.32	2686.049	47	43	TPPVVNPPQVPSHYAPNYPIGHPK*	1PO4
2695.33	2695.067	50	49	KPIDYSMLDEIGHGINSQAHSQVR	
2695.33	2695.069	50	50	KPIDYSMLDEIGHGINSQAHSQVR	
2762.45	2762.163	47	44	TPPVVNPPQVPSHYAPNYPIGHPKR	
2762.45	2762.173	47	45	TPPVVNPPQVPSHYAPNYPIGHPKR	
2775.29	2775.072	50	46	KPIDYSMLDEIGHGINSQAHSQVR	1PO4
2842.42	2842.143	47	43	TPPVVNPPQVPSHYAPNYPIGHPKR*	1PO4
2842.42	2842.155	47	44	TPPVVNPPQVPSHYAPNYPIGHPKR*	1PO4
2993.22	2993.152	64	68	NDDGWWEGVMDGVTGLFPGNYVEPCV	1PO4
3054.56	3054.2	47	49	EYRTPPVVNPPQVPSHYAPNYPIGHPK	
3075.62	3075.287	56	53	TSTGSQLAPIVPEDQNLPGWVWPKNFIEK	
3121.61	3121.31	46	46	GSSHGSVQSLLPPSVGPPPTTKPPTPPQMSR	
3121.61	3121.281	46	47	GSSHGSVQSLLPPSVGPPPTTKPPTPPQMSR	
3172.56	3172.315	62	58	VVAIYDYYADKDELSFQESSVLYVLK	
3201.58	3201.275	46	46	GSSHGSVQSLLPPSVGPPPTTKPPTPPQMSR*	1PO4
3201.58	3201.294	46	47	GSSHGSVQSLLPPSVGPPPTTKPPTPPQMSR*	1PO4
3340.59	3340.264	50	48	SSMPPAPPSPLTVSQHEMTEQSHIGMHTLGR	

*Confirmed ABI phosphopeptides determined within the study are outlined in red. All PTM candidate masses were validated by MS/MS and manual interpretation.

MS/MS inclusion lists based on ABI sequence, *m/z* and theoretical retention times were used as the primary criteria for precursor ion selection (Table 4.6). Ions falling within 0.4 Da and +/- 4 minutes of predicted LC retention time were identified in MS mode, including candidate ions *m/z* 2686 (Fraction 43, Figure 4.9) and *m/z* 3201 (Fraction 46, Figure 4.10). Analysis of fragmentation patterns led to the identification of previously uncharacterized ABI phosphopeptides TPPVNPQVPSHpYAPNYPIGHPK (Figure 4.9A) and GSSHGSVQSLLPPSVGPPPTTKPPpTPPQMSR (Figure 4.9B).

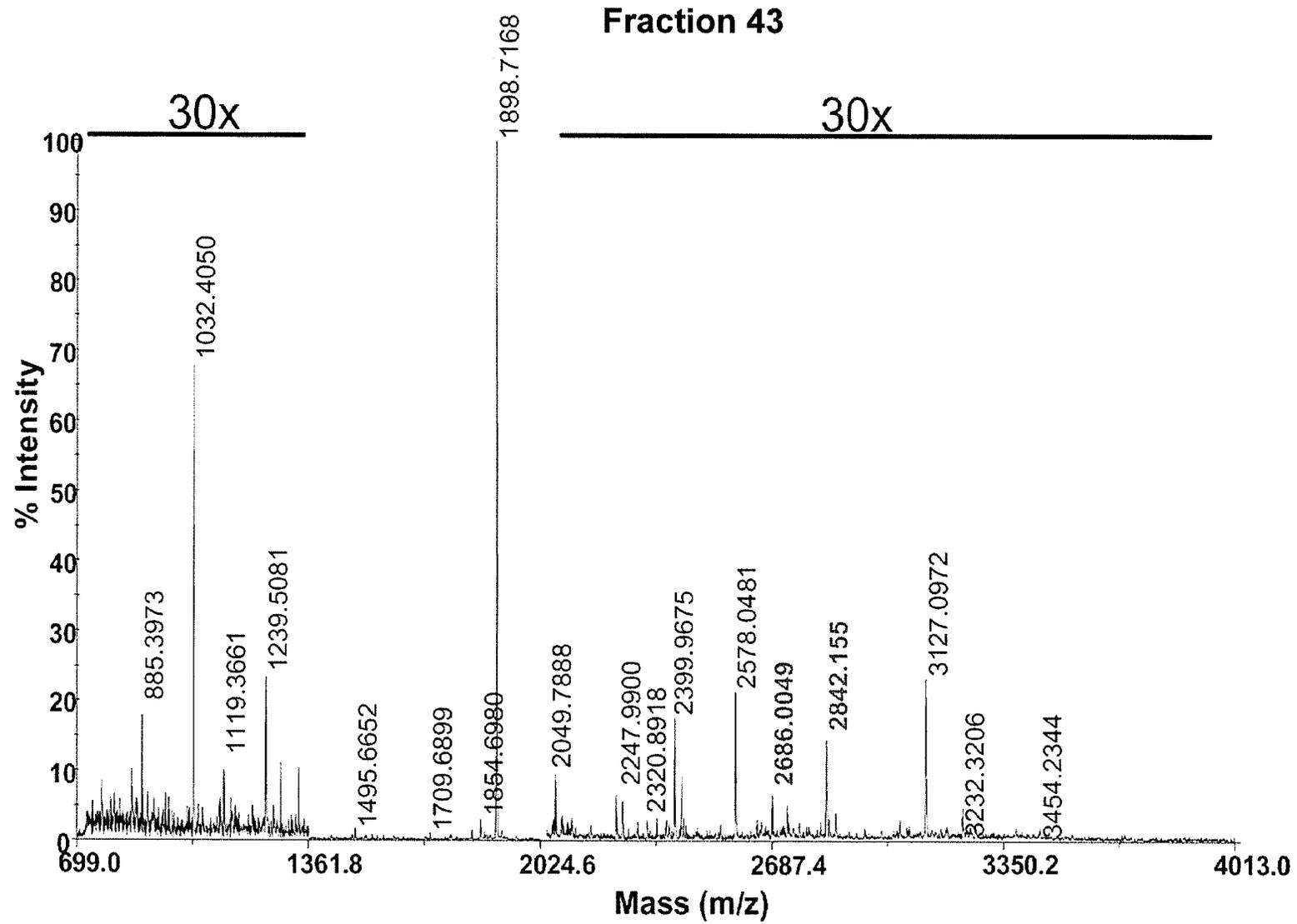


Figure 4.8: Off-line HPLC-MALDI-MS of Fraction 43. M/z of phosphopeptide candidates within the ABI MS/MS inclusion list (see Table 5.6) are indicated in red.

Fraction 46 15x

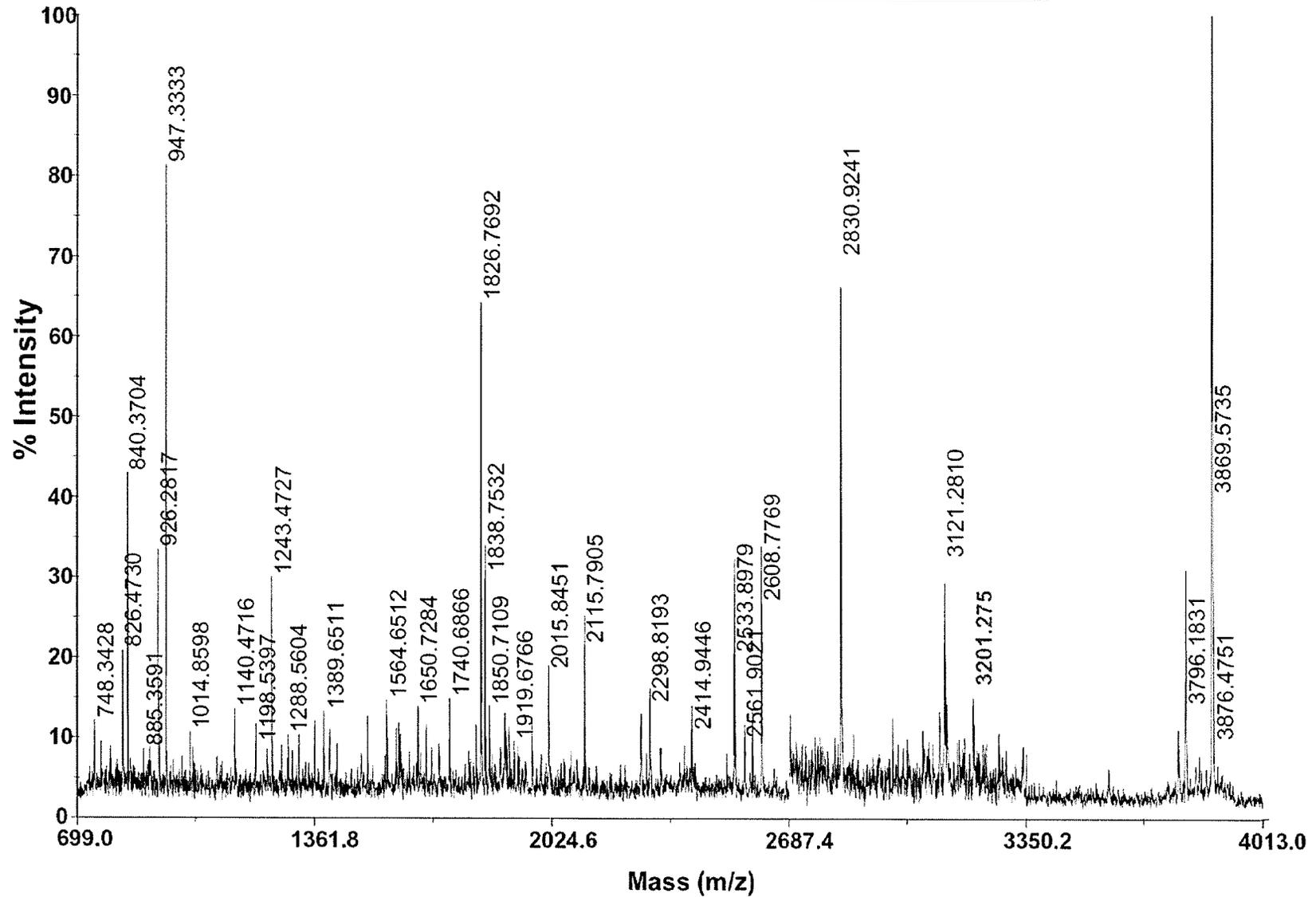


Figure 4.9: Off-line HPLC-MALDI-MS of Fraction 46. M/z of phosphopeptide candidate within the ABI MS/MS inclusion list (see Table 5.6) indicated in red.

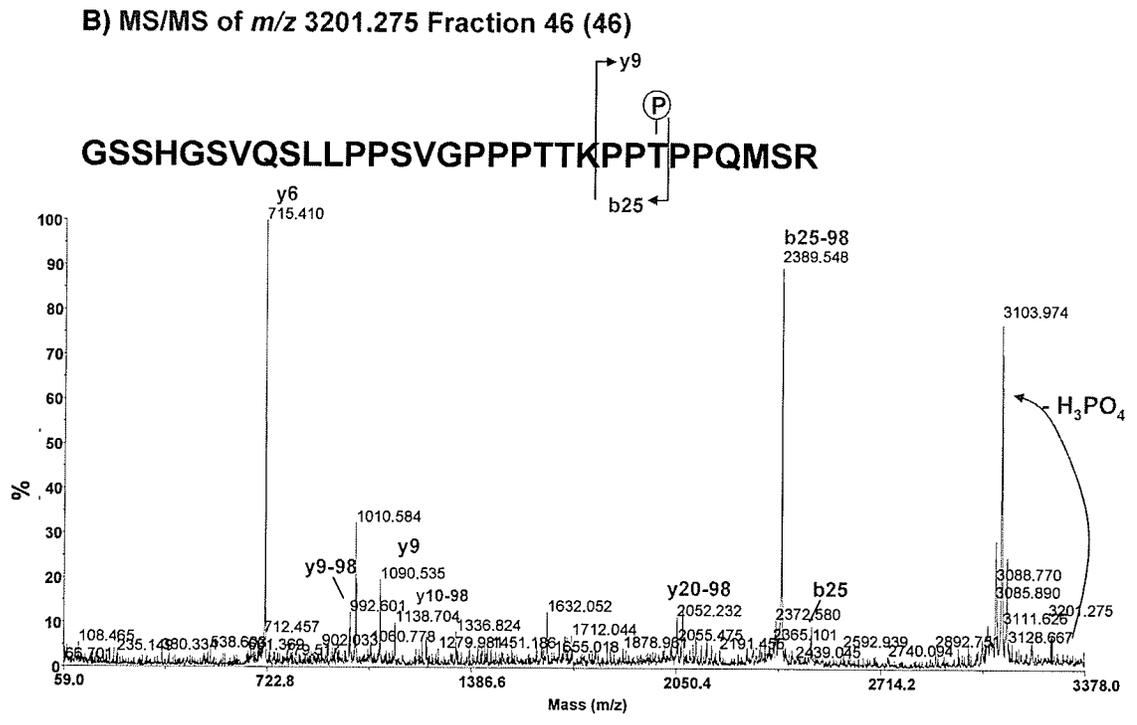
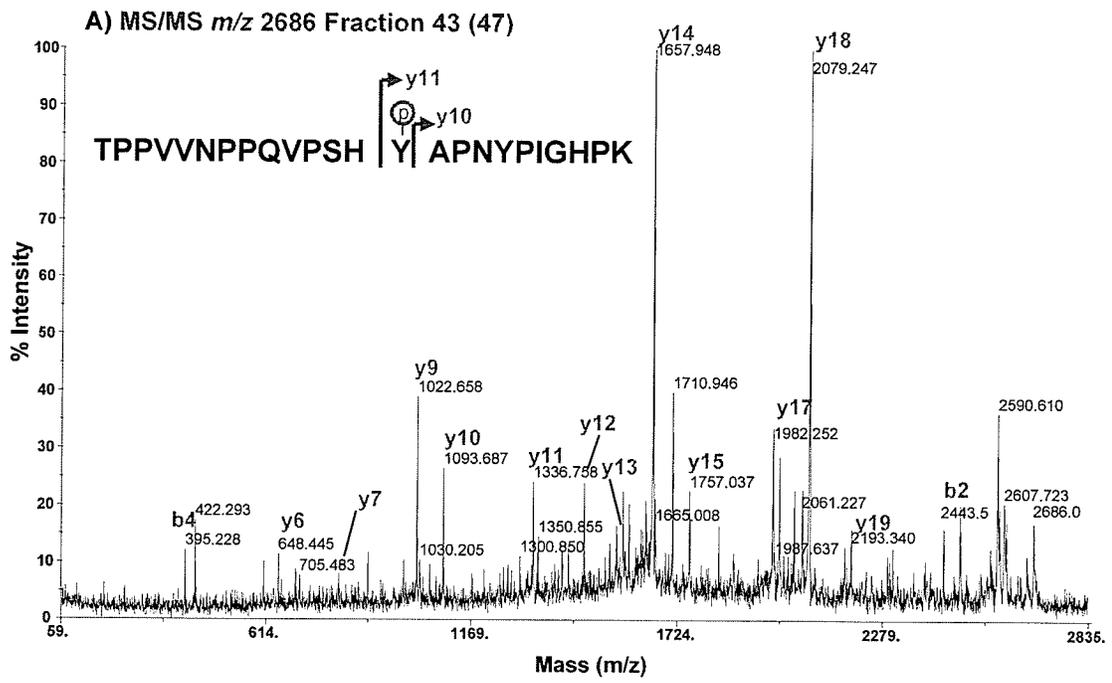


Figure 4.10: MS/MS spectra of ABI phosphopeptides identified using inclusion lists based on LC retention time and m/z . MALDI-MS/MS spectra of

(A) m/z 2686.049 and (B) 3201.275. Ion series assignments and the identification of ABI protein phosphorylation sites (inset). Experimentally observed fraction and theoretical retention time fraction numbers (brackets) are indicated.

4.4 Results and Discussion

The maturation of proteomic technologies and the development of phosphopeptide enrichment strategies have made mass spectrometry the method of choice for many proteomic efforts targeting cell-cycle regulation, growth and intercellular communication. Proteomic researchers are often confronted with the necessity of identifying protein phosphorylation and/or tracking phosphorylation dynamics in a hypothesis-driven manner. The indiscriminant acquisition of MS/MS data based on ion intensity, however, does not properly address bias towards the identification of high abundance peptides. Due to these limitations we have undertaken a targeted approach for the identification of phosphopeptides based on both m/z and phosphopeptide LC retention time. Consistent with the introduction of negative charge/ionic character from the addition of a phosphate group, an overall negative mean phosphopeptide LC retention time deviation was observed and summarized in Figure 4.11. While studies presented here are primarily proof-of-concept, data suggest that the use of LC retention time prediction is a valuable and helpful constraint for the identification of phosphopeptides within a set of LC deposited sample spots.

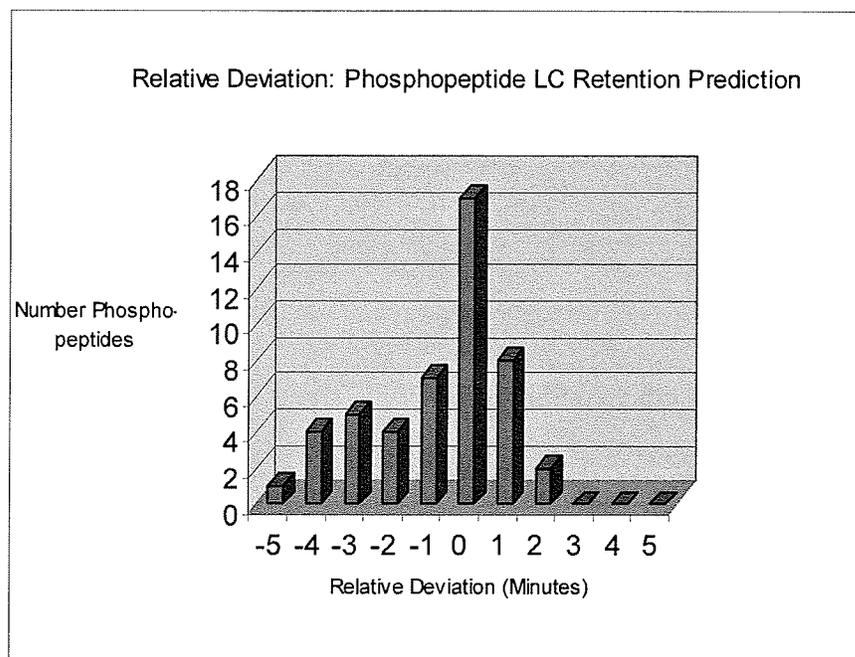


Figure 4.11: Relative deviations of experimentally confirmed phosphopeptides and retention times compared to predicted values. In total, retention time values for 19 distinct phosphopeptides over multiple runs were measured and presented here. In general, 98%, 91%, 81%, 68% and 36% of phosphopeptides eluted within +/- 4, 3, 2, 1 and 0 minutes of predicted LC retention time respectively.

4.4.1 Seven-protein digest mixture.

Inclusion lists based on phosphopeptide LC retention time and m/z were first tested using a protein mixture incorporating phosphorylated α and β -caseins standards. Tryptic peptides spanning a maximum of two tryptic cleavage sites were used to generate an experimental pool of casein phosphopeptides (Table 4.1). Replicate LC separations were performed (Table 4.2) to account for possible run-to-run variations and to fine-tune the method for subsequent studies. A total of 7 phosphorylated α - and β -casein peptides were identified (including TVDMEpSTEVFTK, TVDEMpSTEVFTKK, VPQLEIVPNpSAEER, YLGEYLIVPNpSAEER, DIGpSEpSTEDQAMEDIK, YKVPQLEIVPNpSAEER, FQpSEEQQQTEDELQDK). Peptide measurements presented here were all confirmed by MS/MS sequencing, including phosphopeptides FQpSEEQQQTEDELQDK (m/z 2061, fraction 36, predicted fraction 36, Figure 4.4A) and VPQLEIVPNpSAEE (m/z 1660, fraction 42, predicted fraction 42, Figure 4.4B). As all phosphopeptides fell within \pm 3 minutes (with the vast majority correctly predicted, i.e. \pm 0 minutes) of expected, the study was extended to include the analysis of phosphorylation sites within PKA phosphorylated bovine myelin basic protein (MBP).

4.4.2 Myelin basic protein (MPB).

Myelin basic protein (MBP) is a principal component of the central nervous system, contributing approximately 30% of total protein mass, which is thought to undergoes rapid phosphorylation and dephosphorylation *in vivo*

(Boggs et al. 2006). Previous studies by Hsiao et al., 2007 indicated that protein kinase A treatment of MBP would be a suitable method to generate an additional set of phosphopeptides for method evaluation. *In vitro* phosphorylated MBP was subjected to tryptic digestion and HPLC-MALDI-MS analysis. LC retention times were calculated for theoretical MBP tryptic fragments, having consideration for partial phosphorylation at S/T/Y residues. *In silico* analysis of MS data identified ten phosphopeptide for subsequent MS/MS analysis (Table 4.3) including HRDpTGILDSGR (m/z 1419.7, fraction 37, predicted fraction 40) and NIVTPRpTPPSQGK (m/z 1571.8, fraction 45, predicted fraction 49) presented in Figure 4.5. Inter-run comparisons suggested that the workflow was capable of determining phosphopeptide elution times within a standard deviation of 0.6 to 2.6 minutes and a relative mean accuracy of -4.3% to 1.9 % as presented in Table 4.3.

4.4.3 Anti-pY Immunoprecipitates: Abelson-interacting protein (ABI).

To further demonstrate the utility of the method, cell lysates of *Drosophila* S2 cells were analyzed. These cells are the subject of biological assay within the Institute for Biological Chemistry (Academia Sinica, Taipei, Taiwan) and because of the availability of this lysate it was selected to evaluate the analytical workflow and represent a 'real world' sample. To reduce complexity and enrich the phosphoprotein component within the sample, a Sigma monoclonal anti-pY antibody was used for immunoprecipitation. Proteins isolated were selectively released with

phenylphosphate, digested and separated and analyzed by off-line HPLC-MALDI-MS/MS. Intensity-driven data-dependent analysis led to the identification of more than 60 proteins including Abelson-interacting protein (ABI), daughter of sevenless (DOS), Fak-like tyrosine kinase, receptor tyrosine phosphatase and PDGF/VEGF receptor tyrosine kinase (Supplementary information, Section 7.0). Despite the acquisition of more than 530 MS/MS spectra and inclusion of partial S/T/Y phosphorylation into database searches, only one phosphopeptide was confidently identified within the LC run. ABI peptide TPPVNNPPQVPSHYAPNYPIGHPKR + Phospho Y was identified (Table 4.4) and validated (Figure 4.6). Because of its demonstrated phosphorylation and known biological importance as a regulator and substrate of Abelson protein tyrosine kinase (Wang et al. 1996; Mayer and Baltimore, 1994), the ABI sequence became the subject for LC retention time prediction using the inclusion list strategy.

Generation of the MS/MS inclusion list was aided by applying *Search Peak* program and analyzing MS data, ABI sequences, *m/z* and phosphopeptide LC elution time led to the identification of several candidate ions summarized in Table 4.6. All suspected phosphopeptide ions were subjected to MS/MS fragmentation which led to the identification of two new ABI phosphopeptides not previously identified within automated acquisitions (Figure 4.10). Subsequent MS/MS analysis confirmed the identities of all ABI phosphopeptides namely TPPVNNPPQVPSpYAPNYPIGHPK (*m/z* 2686,

fraction 43, predicted fraction 47), TPPVVNPPQVPSpYAPNYPIGHPKR (*m/z* 2842, fraction 43, predicted fraction 47) and GSSHGSVQSLPPSVGPPPTTKPPpTPPQMSR (*m/z* 3201, fraction 46, predicted fraction 46). Despite having concentrations within instrumental dynamic range, MS intensities *m/z* 2686 and *m/z* 3201 did not trigger fragmentation attempts within the initial proteomic survey presumably due to their low MS intensities (Figures 4.8 and 4.9).

Over the course of the study, use of the inclusion list strategy led to a reduced sample consumption and increased sensitivity towards low abundant phosphopeptides, compared with the indiscriminant 'top 5' ion intensity MS/MS trigger function. Although the inclusion of 'top 10' or '15' intensities into automated runs would theoretically yield more phosphopeptide identities, our experience indicates that this approach merely led to higher rates of sample consumption and larger numbers of poor quality spectra, and a reduction in analytical sensitivity. For example, although automation was capable of generating significantly more MS/MS spectra ($N = 536$) for protein identification, the application of the inclusion list strategy resulted in a 70 fold ($N = 14$, 14% vs. 0.19%) enhancement of identification efficiency (# of phosphopeptide identifications per spectra acquired). While it can be argued that automation did enhance phosphopeptide detection (by providing protein identities), future workflows incorporating both automated and manual workflows could be easily performed to interrogate a single LC-MALDI sample

set (use of Top 1 or 2 intensities, followed by phosphopeptide retention time evaluation).

Although manual efforts were required, we propose that *m/z* and LC retention time is a suitable constraint to increase the number of phosphopeptide specific fragmentation attempts. Using automated runs as an example, biases towards high abundance peptides are clearly demonstrated by the repeated acquisitions of hrp48.1 (Supplementary Information, rank 1, MASCOT score 1648,) and SCAR (Supplementary Information, rank 2, MASCOT score 1074) peptides (Table 4.4). More specifically, peptides AFDQQIFSR (SCAR, 4x), AQAWATGGPSTTGPGVGGMPR (hrp48.1, 3x) and QEGASNYGAGPR (hrp48.1, 3x), while GFGFVTFADPTNVNHVLQNGPHTLDGR (hrp48.1, 2x) and TIDPKPCNPR (hrp48.1, 2x) were each subjected to multiple fragmentation attempts. In other words, although 4 peptides were identified, the relatively large quantities of MS/MS data represents unnecessary sample consumption, redundant data, instrumentation usage and processing time. We have shown that the use of inclusion lists based on *m/z* and predicted LC retention time is effective in increasing the number, quality, and sensitivity of phosphopeptide detection. Although a number of the procedures described in this study were conducted manually, integrative steps to automate the workflow should be explored. Similarly, 'exclusion lists' based on similar principles would also reduce the

repeated detection of high abundance proteins/contaminant and enhance phosphopeptide sensitivity.

4.5 Acknowledgements

The authors would like to gratefully acknowledge support (to V.C.) from the National Science and Engineering Council (NSERC, Canada), the Science and Technology Division of the Taipei Economic and Cultural Office (TECO, Canada), and the National Science Council (NSC, Taiwan) for international travel and collaborative research funding. Dr. T.C. Meng (Institute for Biological Chemistry, Academia Sinica) is gratefully acknowledged for providing S2 cells. This work was also supported by a Genomic Medicine Grant from the National Science Council (NSC 91-3112-P-001-009-Y) to the National Core Facilities for Proteomics Research, located at the Institute of Biological Chemistry, Academia Sinica, Taiwan.

4.6 References

- Chen, E.I.; Hewel, J.; Felding-Habermann, B.; Yates, J.R. (2006). Large scale protein identification by combination of protein fractionation and multidimensional protein identification technology (MudPIT). *Mol Cell Proteomic.* 5, 53-56.
- Chen, V.C., Cheng, K., Ens, W., Standing, K. G., Nagy, J. I., Perreault, H. (2004) Device for the reversed-phase separation and on-target deposition of peptides incorporating a hydrophobic sample barrier for matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 76, 1189-1196.
- Ens, W.; Standing, K.G. (2005). Hybrid quadruple/time-of-flight mass spectrometers for analysis of biomolecules. *Methods Enzymol.* 402, 49-78.
- Hunter, T. (2000) Signaling-2000 and beyond. *Cell* 100, 113-27
- Karas, M.; Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 Daltons. *Anal Chem.* 60, 2299-301.
- Krokhin. O.V.; Craig, R.; Spicer, V.; Ens ,W.; Standing, K.G.; Beavis, R.C.; Wilkins, J.A. (2004) An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. *Mol. Cell Proteomics.* 3, 908-19.
- Krokhin. O.V. (2006)^a Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300- and 100-Å pore size C18 sorbents. *Anal. Chem.* 78, 7785-95.
- Krokhin. O.V. (2006)^b Use of peptide retention time prediction for protein identification by off-line reversed-phase HPLC-MALDI MS/MS. *Anal. Chem.* 78, 6265-9.
- Krokhin, O.V.; Ens, W.; Standing, K.G. (2005) MALDI QqTOF MS combined with off-line HPLC for characterization of protein primary structure and post-translational modifications. *J. Biomol. Tech.* 16, 429-40.
- Lee, C. L.; Hsiao, H. H.; Lin, C. W.; Wu, S. P.; Huang, S. Y.; Wu, C. Y.; Wang, A. H.; Khoo, K. H., (2003) Strategic shotgun proteomics approach for efficient construction of an expression map of targeted protein families in hepatoma cell lines. *Proteomics* 3, 2472-86.
- Lin, T.Y.; Huang, C.H.; Chou, W.G.; Juang, J.L. (2004) Abi enhances Abl-mediated CDC2 phosphorylation and inactivation. *J. Biomed. Sci.* 11, 902-910.

- Pawson, T. (2007). Dynamic control of signaling by modular adaptor proteins. *Curr Opin Cell Biol.* 19, 112-116.
- Pawson, T.; Scott, J.D. (2005) Protein phosphorylation in signaling-50 years and counting. *Trends Biochem. Sci.* 30,286-290.
- Petritis, K.; Kangas L. J.; Ferguson, P. L.; Anderson, G. A.; Pasa-Tolic, L.; Lipton, M. S.; Auberry, K. J.; Strittmatter, E. F.; Shen, Y.; Zhao, R.; Smith, R. D. (2003) Use of artificial neural networks for the accurate prediction of peptide liquid chromatography elution times in proteome analyses. *Anal. Chem.* 75, 1039-48.
- Petritis, K.; Kangas; L. J.; Yan, B.; Monroe, M. E.; Strittmatter, E. F.; Qian W. J.; Adkins, J. N.; Moore, R. J.; Xu, Y.; Lipton, M. S.; Camp, D. G. 2nd; Smith, R. D. (2006) Improved peptide elution time prediction for reversed-phase liquid chromatography-MS by incorporating peptide sequence information. *Anal. Chem.* 78, 5026-39.
- Picotti, P.; Aebersold, R.; Domon, B. (2007). The implications of proteolytic background for shotgun proteomics. *Mol. Cell Proteomics. MCP online* 10.1074/mcp.M700029-MCP200.
- Tripet, B.; Cepeniene, D.; Kovacs, J.M.; Mant, C.T.; Krokhin, O.V.; Hodges, R.S. (2006) Requirements for prediction of peptide retention time in reversed-phase high-performance liquid chromatography: hydrophilicity/hydrophobicity of side-chains at the N- and C-termini of peptides are dramatically affected by the end-groups and location. *J. Chromatogr A.* 1141, 212-25.
- Wu, S.L.; Kim, J.; Bandle, R.W.; Liotta, L.; Petricoin, E.; Karger, B.L. (2006) Dynamic profiling of the post-translational modifications and interaction partners of epidermal growth factor receptor signaling after stimulation by epidermal growth factor using Extended Range Proteomic Analysis (ERPA). *Mol. Cell Proteomics.* 5, 1610-27.
- Vestal, M.L.; Campbell, J.M. (2005). Tandem Time-of-flight Mass Spectrometry. *Methods Enzymol.* 402, 79-108.

Part 5:
Conclusions and Future Work

5.1: Conclusions

Proteomics is a rapidly evolving field that has demonstrated an enormous potential to significantly impact biological research. Significant strides have been made in the ability to analyze proteins by mass spectrometry with improving sensitivity and throughput. However as a universal technique for proteomic sampling has yet to be discovered, most of the work within the field has been directed towards the development and testing of new analytical methods. Despite enormous efforts, only recently have novel biological systems been discovered using proteomic methods. While the capabilities of high-throughputs are impressive, the indiscriminant generation of large bodies of data is not a substitute for 'good science' unless this data applied in a meaningful way. This view of proteomics does not imply that high-content and high-throughput proteomics are exclusive of one another. It does mean, however, that imaginative applications of proteomics incorporating both hypothesis-driven and hypothesis-generating workflows are needed before the full potential of proteomics is reached, and proteins are placed within their functional contexts. In this work are the design, development and evaluation of novel analytical techniques for the identification of proteins, protein phosphorylation and protein-protein interactions.

5.2 Off-line vacuum-LC and MALDI deposition techniques for MS and MS/MS

Proteomics aims to analyze the total complement of proteins, including post-translational modifications and interaction partners within a cell or tissue type. Although conceptually simple, the characterization of the proteome is far from a routine endeavor as the basic monitoring of proteins has remained a challenging task. Due to the broad dynamic range and variability of gene products, current techniques are unable to explore all of the phenomena that occur at the protein level. For all the advances within the field, little ground has been gained in the development of robust microvolume approaches to one of the most frequently performed tasks within the life sciences – the purification and enrichment of proteins and peptides. Unlike DNA and RNA, there are no bioenzymatic techniques, analogous to PCR that permit the exponential amplification and purification of target proteins. Because of these limitations, protein samples tend to be precious in nature and the development of any technique that provides the high-performance purification of proteins at the smallest scales is of tremendous benefit.

With the proliferation of MS facilities many biologists now have access to a service in which they can submit a sample and are handed back a list of proteins that have been identified by MS. While proteomic centers will generally offer LC-MS evaluations, these services are primarily restricted to on-line LC ESI-MS studies to decrease user-time and increase sample

throughput. While this arrangement does work when the identities of higher abundance proteins are required, problems arise when dealing with samples of clinical value or precious in nature. Although researchers may be interested in analyzing the proteome at higher levels of sensitivity offered by the MALDI-MS protocol, access to specialized HPLCs and deposition robots is generally in short supply due to their high cost of acquisition. In fact, very little attention has been paid to the development of alternative strategies for the off-line separation of peptides for MALDI-MS. In Part 2, a new cost effective technique utilizing features offered by vacuum LC was presented. Eliminating the need for complex separation systems, the method has been directly applied on any digested sample obtained from PAGE or solution. Using a Cx43 peptide spiked with an overwhelming amount of angiotensin II peptide, SepDep purification led to the identification of all amino acid sequences and positive results suggested the approach would be sensitive enough for most proteomic-level studies. Further, tryptic digests of *E.coli* citrate synthase were used as a test protein leading to a peptide sequence coverage greater than 76%. To extend the evaluation, tryptic peptides of β -casein were separated by the SepDep protocol in conjunction with MALDI-TOF-MS and MALDI-QqTOF-MS/MS analysis. This study not only led to the identification of all predicted phosphorylation sites, but also uncovered peptide isobaric sequences and the presence of (unpredictable) non-specific proteolytic products. Completion of the study and refinement of the workflow indicated that the method could be used to separate samples of higher complexity. Since its initial development,

the device has proven to be a valuable tool for the rapid and sensitive identification of peptides and PTMs and subsequent demonstrations have shown that HPLC-quality separations can be obtained on high femto-mole levels of material (Lattova et al., 2007; Chen et al., 2006; Snovida et al., 2006^a; Snovida et al., 2006^b).

With these developments, researchers now are offered the capability to identify multiple proteins from a single sample or SDS-PAGE band without incurring expensive equipment or ongoing costs. For example, gel slices from SDS-PAGE are the most common sample format submitted to proteomic core facilities for protein identification. Instead of relying upon PMF to establish the identities of high abundance peptides using Ziptips or similar SPE techniques, the newly developed device permits the study of low abundance peptides that would normally be difficult to detect without separation. This feature of the SDS-PAGE combined with the SepDep device not only allows for a broad range of proteins to be studied (e.g. use of SDS permits isolation of membrane proteins), but also offers a robust and readily obtainable technique for the multidimensional separation of proteomic samples.

Despite these advances, in its current form the SepDep methodology is not suitable for high-throughput proteomic applications. Recently however, an automated method for the off-line deposition of LC eluent at subatmospheric pressure using a vacuum deposition interface for MALDI-MS

was described (Chen et al., 2005). While this study demonstrated rapid high-resolution separations could be achieved using monolithic columns, a nano-flow HPLC with a robotic deposition interface was ultimately used which negated any cost benefits offered by a vacuum source. While the method described in Part 2 is suitable for low-throughput applications, processing large batches of samples at this time is only possible in an automated setting. Features identified within the original device - namely the use of a vacuum chamber to activate solvent flow and sample deposition deserves further development. This new embodiment (a.k.a. SepDep II; Appendix Part 6) utilizes a small aluminum probe to interface a small sealable gasket to any MALDI surface, multi-well plate or microcentrifuge tube for off-line sample collection. The modular nature of the SepDep II allows sample delivery in an automated setting. While solvent compositions with higher organic content led to higher rates of sampling failure (due to lower surface tensions and sample dispersion), physical contact between the probe and MALDI surface did not interfere with sample deposition using aqueous concentrations of ACN of 60% or less. Despite these limitations, development of this new device is significant as new avenues for proteomic sample processing are offered. For example, outside the realm of reverse-phase liquid chromatography, the most notable application identified is related to the detection of protein-protein interactions/PTM by open tubular affinity chromatography column. Moreover, because the SepDep devices offer the capability to collect LC fractions into any sampling container or surface, the technique possesses a high level of

compatibility with many other downstream analytical technologies including MS, 1D/2D SDS-PAGE and structural biology techniques such as NMR and x-ray crystallography. Recently, two publications described the derivatization of the inner surface of fused silica tubes with fully functional proteins (Hanna et al., 2006; Bakry et al., 2006). While the open tube concept is novel, functionalized columns with reduced backpressures are ideal devices for vacuum-driven technologies such as the SepDep II and the isolation and enrichment of protein-protein interaction complexes. As functional groups are attached to inner surfaces, these columns lack flow restricting resin/matrix beads, frits and unswept spaces. Furthermore because sample, wash and elution solutions are allowed to directly enter and leave the columns, samples could be rapidly processed. Development of affinity approaches such as these will offer significant new methods for the isolation and purification of protein complexes.

5.3 Development of proteomic approaches for the identification and characterization of protein-protein interactions modulated by PDZ domains

An increasing body of evidence indicates that many biochemical processes are facilitated by the establishment of protein-interaction networks in a time- and location-dependent manner. Because of their importance, the study of protein-protein interactions offers insights into numerous cellular processes. The primary difficulty facing protein biochemists attempting to

study protein-protein interactions lies in the challenge of isolating and characterizing multiprotein complexes from whole-cell lysates. Keeping this in mind we chose to evaluate a complex network of proteins associated with a member of the membrane guanylate kinase (MAGUK) family of proteins, zonula occludens-1 (ZO-1). ZO-1 is a multi-PDZ domain protein which sits upon the crossroads of cytoskeletal scaffolding and transcriptional regulation (Balda et al., 2000; Fanning et al., 2002). Due to the importance of ZO-1 and PDZ-domains in general, we have developed an approach for the isolation and characterization of PDZ protein-protein interactions. Recombinant GST-PDZs of ZO-1 were used to pull down protein material within crude cell lysates, and affinity bound proteins were selectively released using a ligand corresponding to the C-terminal tail of Cx36 (a known ZO-1 PDZ-binding protein). Application of the peptide elution strategy not only led to a decrease in non-specific background and recombinant protein, but also an overall reduction of sample complexity within proteomic samples. This simplification of the proteomic sample permitted analysis of well defined SDS-PAGE protein bands for extraction and PMF identification.

As with any proteomic “pull-down” approach, the validity of the putative interaction required further examination to ensure that nonspecific protein aggregates (affinity ligand and column components) were not measured. To aid this process, identified proteins were screened for the presence of C-terminal PDZ binding motifs to establish first order protein-protein

interactions. In general, the convergence of experimental data with known functions of ZO-1 provided reassurance that the workflow was successful. More specifically, proteomic analysis revealed a number of proteins implicated in the cytoplasmic scaffolding of the plasma membrane, including several peptides originating from vimentin, annexin, plectin, spectrin and a member of the α -actinin family. However because of the close sequence similarity between the non-muscle α -actinins (Figure 4.6) and the nature of the PMF process, identification of which α -actinin was present could not be initially and confidently made. To help identify protein specific peptides, the aforementioned SepDep device was used to chromatographically separate and deposit peptides for MS/MS sequencing – which confirmed the presence of α -actinin-4. Subsequent studies demonstrated the physiological relevance of the newly discovered association using a variety of cell and tissue types. From this standpoint, the discovery that α -actinin-4 binds to ZO-1 is exciting and of broad interest to cell biologists and may provide important insights into possible modes of mitogenic signaling.

Although tight junctions, adherens junctions and gap junctions are distinct intercellular junctions, several of their components, including ZO-1, can interact with each other. While an increasing body of evidence has suggested these regions of the cell have an important role in regulating gene expression and cell proliferation (Takahashi et al. 1998; Mesnil et al., 2005; Balda et al., 2000; Mori et al., 2006; Matter et al., 2003), the exact signaling

mechanisms involving these structures has yet to be fully determined. While it is known that ZO-1 is involved with regulation of the ErbB2 gene (Balda et al., 2000), expression of the receptor tyrosine kinase is not dependent on nuclear accumulation of ZO-1 but rather upon the cytoplasmic sequestering of a Y-box transcription factor to the plasma membrane. Interestingly this transcription factor, termed ZO-1-associated nucleic-acid binding protein (ZONAB; Balda et al., 2000), has also been shown to form a complex with a cell-cycle control protein cyclin-dependent kinase 4 (CDK4, Balda et al., 2003).

α -Actinin-4 is a protein found in the majority of cells and tissue types. It is well established that α -actinin-4 is highly concentrated at a variety of junctions, cell-matrix contacts and plasma membrane receptors (Otey et al., 2004). Based on the association discovered in my work, a working model has been proposed in which α -actinin-4 is used to monitor the assembly state of cell-cell and cell-matrix contacts and the regulation of cell proliferation. In this model, under high cell density conditions, the inhibition of ZONAB and CDK4 is achieved by sequestration of α -actinin-4 and ZO-1 to sites of contact within the plasma membrane. Conversely in subconfluent cells, cell-cell/cell-matrix contacts are depleted and inhibitory pressures limiting cell proliferation are removed (Figure 5.1). In this manner, the model suggests α -actinin-4, ZO-1, ZONAB and CDK4 are a part of a system that is responsible for monitoring cell-density by regulating gene expression and cell-cycle control.

While findings related to the binding of ZO-1 with α -actinin-4 have been presented, a number of questions still require further investigation. Most notably, despite the high degree of sequence homology and consensus C-terminal sequences, the observation that α -actinin-4, but not α -actinin-1, is capable of binding to ZO-1 poses a perplexing question: what sequence differences of α -actinin-4/ α -actinin-1 govern ZO-1 recognition? Structural analysis of the PDZ1 with α -actinin-1/ α -actinin-4 by crystallography would no doubt help to address this question.

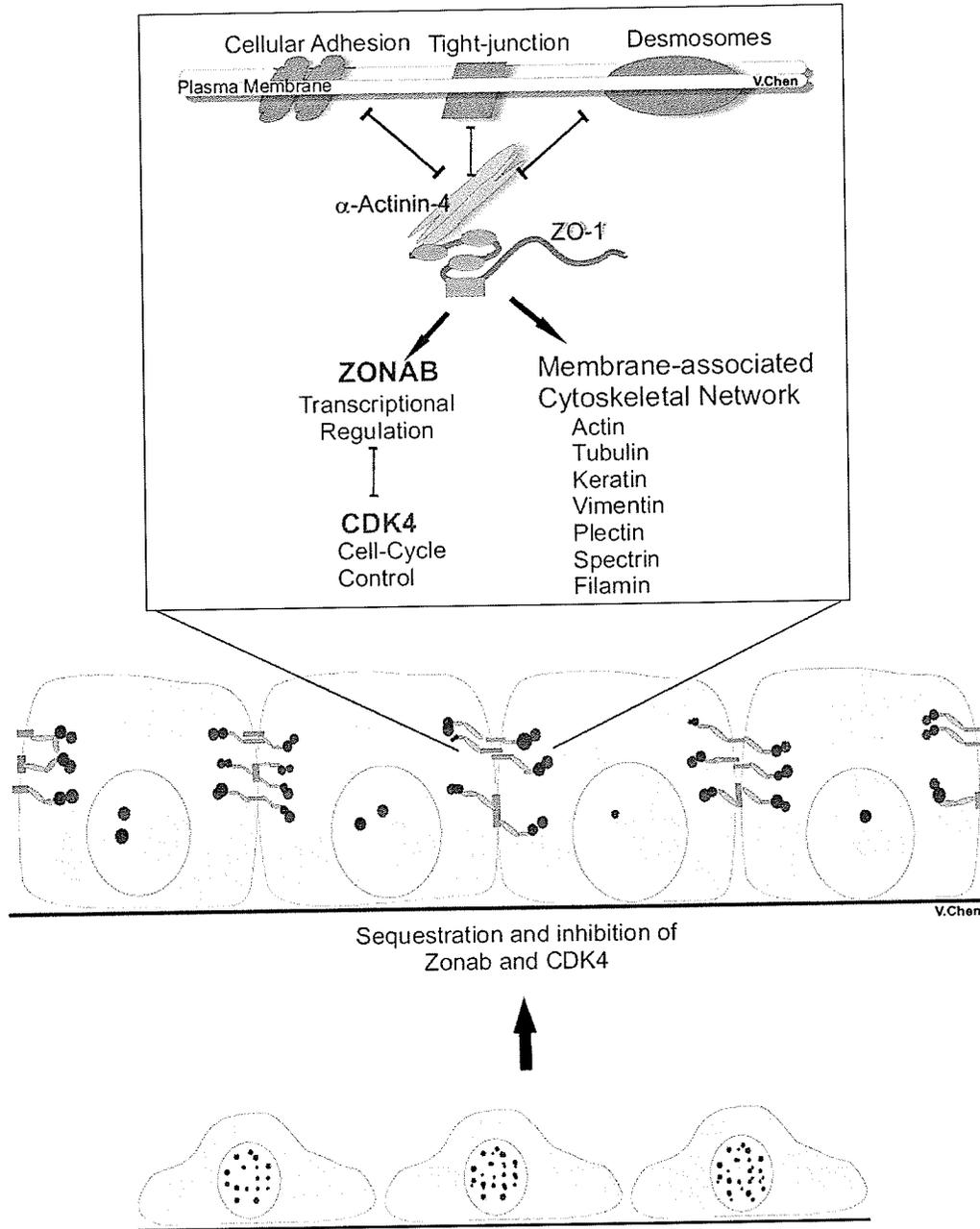


Figure 5.1: ZO-1/ α -actinin-4 mediated transcriptional regulation. The lower figure shows low-density epithelial cells with low levels of cell-cell and cell-matrix contacts. The accumulation of α -actinin-4 at TJ, FA and AJ with increasing cell density may not only function to link the cytoskeleton with the plasma membrane, but also serve to sense cell density by recruiting α -actinin-4 (yellow) and thereby scaffolding ZO-1 (shown in light blue), ZONAB and CDK4 (represented as red and blue circles).

As the PDZ1 pull-down approach was successful in the identification of several proteins linked to the ZO-1 complex, subsequent studies should include the removal of sequential slices of the SDS-PAGE (instead of removing individual bands) to collectively represent all the protein. Another limitation is related to the use of PMF as the principal identification strategy. While PMF is successful in identifying proteins of high abundance, application of LC-MALDI would dramatically increase the chance of identifying important proteins of low abundance. Further, proteomic characterizations using the GST-PDZ2 and GST-PDZ3 along with suitable peptides for selective release of protein interactions, i.e. C-terminal tails of Cx43 (PDZ2, Giepmans et al. 1998) and PLC β 3 (PDZ3, Van Zeijl et al., 2007) would also be of tremendous value to establish the array of proteins linked to ZO-1. Finally, the application of SILAC (described in section 1.2.2, Figure 1.2, Blagoev et al., 2003) should also be considered (as described by Figure 5.2) and would no doubt be beneficial to by-pass the problems associated with the detection of non-specific protein.

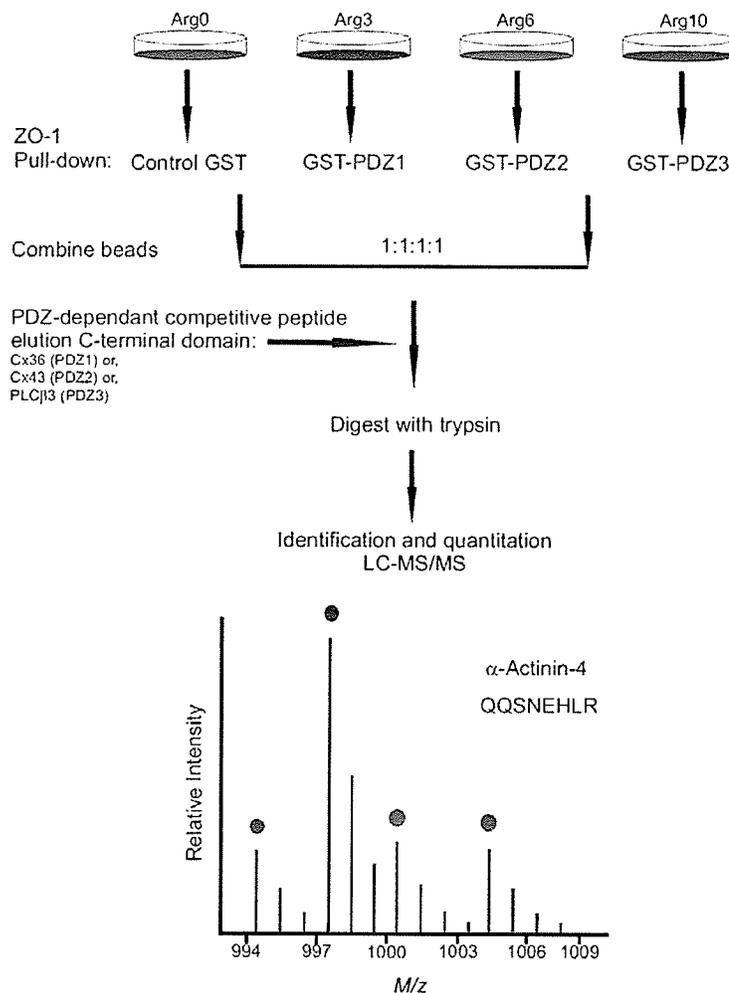


Figure 5.2: Identification of protein-protein interactions of ZO-1 PDZ1, PDZ2 and PDZ3 domains by SILAC. Stable isotopically labeled cell cultures (SILAC, using native and isotopic labeled arginine: Arg0, Arg3, Arg6 and Arg10) is used to analyze the PDZ1-3 and establish a GST control. After incubation beads are collected, wash and mixed, pull-downs then are incubated with a C-terminal PDZ binding ligand (e.g. of Cx36 (PDZ1, Li et al., 2004), Cx43 (PDZ2, Giepmans et al., 2004) or PLCβ3 (PDZ3, van Zeijl et al., 2007)) to specifically release associated proteins. Peptides demonstrating a specific interaction will exhibit statistically significant ratios. Experiments can also be repeated with a variety of competitive binding ligands to systematically probe protein-protein interactions modulated by each domain.

Part 5.4: Development of off-line HPLC MALDI-MS/MS methods for high-content phosphoproteomics using LC retention time prediction

Phosphorylation is one of the most prevalent and important protein modifications. The modification plays a key role in the regulation of a number of different cellular processes including complex formation, signal transduction, activation/inactivation of enzymes, and protein degradation. The abnormal activation of kinases or signaling pathways by phosphorylation has been implicated in a number of disorders including numerous forms of cancer. The experimental determination of phosphorylation sites is an important, yet challenging task for establishing protein function and their modes of regulation.

Although the introduction of automation has revolutionized proteomic analysis, we have demonstrated that data-dependent MS/MS acquisitions based on ion intensity can significantly lower phosphopeptide detection limits. While many phosphoproteomic studies employ “data-dependent” or “MS/MS triggering” as a function of ion intensity, we have shown a significant number of phosphopeptide signals are likely lost, despite being well within instrumental dynamic range. To increase success rates within phosphoproteomic studies we have demonstrated the use of LC retention time prediction and ion *mass-to-charge* to target specific peptides carrying the post-translational modification. Further, although it is well recognized that phosphopeptides do have similar retention times to their non-modified forms,

this relationship has never been quantified. While the use of the retention time prediction workflow did lead to an increase in phosphopeptide identifications within immunoprecipitates of S2 cell proteins, we have also noted a dramatic reduction of redundant data and lower rates of sample usage. While many of the steps described within Part 4 were manually performed, it should be noted that they could be easily integrated into an automated setting.

5.6 References

- Blagoev, B.; Kratchmarova, I.; Ong, S.E.; Nielsen, M.; Foster, L.J.; Mann, M. (2003). A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nature Biotechnol.* 21, 315-318.
- Chen, H.S.; Rejtar, T.; Andreev, V.; Moskovets, E.; Karger, B.L. (2005). High-speed, high-resolution monolithic capillary LC-MALDI MS using an off-line continuous deposition interface for proteomic analysis. *Anal. Chem.* 77, 2323-31.
- Chen, V.C.; Cheng, K.; Ens, W.; Standing, K.G.; Nagy, J.I.; Perreault, H. (2004). Device for the reversed-phase separation and on-target deposition of peptides incorporating a hydrophobic sample barrier for matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 76, 1189-1196.
- Chen, V.C.; Li, X.; Perreault, H. Nagy, J.I. (2006). Interaction of zonula occludens-1 (ZO-1) with alpha-actinin-4: application of functional proteomics for identification of PDZ domain-associated proteins. *J. Proteome Res.* 5, 2123-2134.
- Giepmans, B.N. (2004) Gap junctions and connexin-interacting proteins. *Cardiovasc. Res.* 62, 233-45.
- Giepmans BN, Verlaan I, Moolenaar WH. (2001) Connexin-43 interactions with ZO-1 and alpha- and beta-tubulin. *Cell Commun. Adhes.* 8, 219-23.
- Harris, B.L.; Lim, W.A. (2001). Mechanism and role of PDZ domains in signaling complex assembly. *J. Cell. Sci.* 114, 3219-31.

- Honda, K., Yamada, T., Endo, R., Ino, Y., Gotoh, M., Tsuda, H., Yamada, Y., Chiba, H., Hirohashi, S. (1998) Actinin-4, a novel actin-bundling protein associated with cell motility and cancer invasion. *J. Cell Biol.* 140, 1383-1393.
- Howarth, A.G. Stevenson, B.R. (1995) Molecular environment of ZO-1 in epithelial and non-epithelial cells. *Cell Motil Cytoskelton.* 31, 323-32.
- Lattova, E.; Chen, V.C.; Varma, S.; Bezabeh, T.; Perreault, H. (2007). Matrix-assisted laser desorption/ionization on-target method for the investigation of oligosaccharides and glycosylation sites in glycopeptides and glycoproteins. *Rapid. Com. Mass Spectrom.* 21, 1644-50.
- Li, X.; Olson, C.; Lu, S.; Kamasawa, N.; Yasumura, T.; Rash, J.E.; Nagy, J.I. (2004). Neuronal connexin36 association with zonula occludens-1 protein (ZO-1) in mouse brain and interaction with the first PDZ domain of ZO-1. *Eur J Neurosci.* 19, 2132-46.
- Matter, K.; Balda, M. (2003). Signalling to and from tight junctions. *Nat. Rev. Mol. Cell Biol.* 4, 225-36.
- Mesnil, M.; Crespin, S; Avanzo, J., and Zaidan-Dagli, M. (2005). Defective gap junctional intercellular communication in the carcinogenic process. *Biochim. Biophys. Acta.* 1719, 125-145.
- Mori, R.; Power, K.T.; Wang, C.M.; Martin, P.; Becker, D.L. (2006) Acute downregulation of connexin43 at wound sites leads to a reduced inflammatory response, enhance keratinocyte proliferation and wound fibroblast migration. *J. Cell Sci.* 119, 5193-5203.
- Otey, C. A., Carpen, O. (2004) Alpha-actinin revisited: a fresh look at an old player. *Cell Motil. Cytoskeleton.* 58, 104-111.
- Reichert, M.; Muller, T. Hunziker, W. (2000) The PDZ domains of zonula occludens-1 induce an epithelial to mesenchymal transition of Madin-Darby canine kidney 1 clls. Evidence for a role of β -catenin/Tcf/Lef signaling. *J. Biol. Chem.* 275, 9492-9500.
- Snovida, S.I.; Chen, V.C.; Perreault, H. (2006^a). Use of a 2,5-dihydroxybenzoic acid/aniline MALDI matrix for improved detection and on-target Derivatization of glycans: A preliminary report. *Anal. Chem.* 78, 8561-8568.
- Snovida, S.I.; Chen, V.C.; Krokhin, O.; Perreault, H. (2006^b). Isolation and identification of sialylated glycopeptides from bovine alpha1-acid

glycoprotein by off-line capillary electrophoresis MALDI-TOF mass spectrometry. *Anal. Chem.* 78, 6556-6563.

Taliana, L., Benezra, M., Greenberg, R.S., Masur, S.K., Bernstein A.M. (2005) ZO-1: Lamellipodial localization in a corneal fibroblast wound model. *Invest. Ophthalmol. Vis. Sci.* 46, 96-103.

Takahashi, K.; Matsuo, T.; Katsube, T.; Ueda, R. and Yamamoto, D. (1998). Direct binding between two PDZ domain proteins Canoe and ZO-1 and their roles in regulation of the jun N-terminal kinase pathway in *Drosophila* morphogenesis. *Mech. Dev.* 78, 97-111.

Torrado, M., Senatorov, V.V., Trivedi, R., Fariss, R.N., Tomarev, S.I. (2004) Pdlim2, a novel PDZ-LIM domain protein, interacts with alpha-actinins and filamin A. *Invest Ophthalmol Vis Sci.* 45, 3955-63.

Van Zeijl, L.; Ponsioen, B.; Giepmans, B.N.G.; Ariaens, A.; Postma, F.R.; Varnai, P.; Ballen, T.; Divecha, N.; Jalink, K.; Moolenaar, W.H. (2007). Regulation of connexins43 gap junctional communication by phosphatidylinositol 4,5-bisphosphate. *J. Cell Biol.* 177, 881-891.

6.0 Appendix

METHOD AND APPARATUS FOR DEPOSITING SAMPLES ON A TARGET SURFACE

METHOD AND APPARATUS FOR DEPOSITING SAMPLES ON A TARGET SURFACE

Vincent C. Chen and H el ene Perreault

Invention description:

Vincent C. Chen and H el ene Perreault "METHOD AND APPARATUS FOR DEPOSITING
SAMPLES ON A TARGET SURFACE" Provisional Patent, 2005*

Vincent C. Chen and H el ene Perreault "METHOD AND APPARATUS FOR DEPOSITING
SAMPLES ON A TARGET SURFACE" United States and Canada PCT Patent, 2007* United
States Application 20070023681

*Written in collaboration with Luc Berube, Berskin and Parr Attorneys and Patent Agents.

Portions of this work maybe considered confidential communication.

6.1 Abstract

The off-line collection of LC effluent onto MALDI sample supports has received interest due to recent advances in MALDI-MS/MS instrumentation. We have demonstrated the utility of negative pressure liquid chromatography (SepDep) to increase sequence coverage in proteomic research. Here we report an automated embodiment of the original SepDep (Figure 2.1), for off-line LC and sample deposition for MALDI-MS (Vincent C. Chen and Hélène Perreault "METHOD AND APPARATUS FOR DEPOSITING SAMPLES ON A TARGET SURFACE" United States and Canada PCT Patent, 2007 United States Application 20070023681). The novel device incorporates an automated x-y-z stage and miniaturized vacuum interface capable of carrying out LC separation and deposition on to any commercially available MALDI target. Design and construction of the interface has been completed with the purpose of evaluating performance characteristics for the purpose of intellectual property protection. Advantageous features identified within the original embodiment were incorporated into the SepDep II design, including mechanisms of vacuum solvent flow generation and MALDI deposition. Evaluation of the interface indicates that the physical size of the apparatus is only limited by the size of the MALDI sample spot desired. This novel design offers an all-in-one device capable of off-line solid phase extraction, capillary electrophoresis or liquid chromatography for MALDI-MS and MS/MS analysis. It is expected the development of this device will offer new avenues for multidimensional separations in proteomic research.

6.2 FIELD OF THE INVENTION

[1] The invention relates to the field of macromolecules separating device and more specifically to the field of macromolecules separating device for mass spectrometry analysis.

6.3 BACKGROUND OF THE INVENTION

[2] Mass spectrometry has become a major analytical tool in proteomic and biological research in general. Most protein identification strategies involving MS analyze proteolytic peptides (e.g. tryptic digests for mass fingerprinting in combination with tandem MS to confirm amino acid sequence and the presence of various post-translational modifications. In most proteome studies, proteins are separated on electrophoretic gels and in-gel digestion extracts are subjected to MS analysis. Although matrix assisted laser desorption/ionization (MALDI) is very effective for screening high abundance proteins in complex samples, lower abundance peptides often remain undetected. Suppression effects are a common problem arising from the presence of multiple analytes competing for protons during the ionization process. Various separation methods have been coupled to mass spectrometry to improve identification of macromolecules. However, efficient coupling to MALDI has been more difficult. High throughput MALDI requires deposition of multiple sample drops on a MALDI target plate. Chen et al. (*Analytical Chemistry* vol. 76, No 4, 2004) have proposed a method for separating molecules within samples and depositing drops of the eluent on a

MALDI target surface by using negative pressure. However, their apparatus comprises a cumbersome vacuum chamber box in which the MALDI target and the separation device are inserted. Drop deposition is made difficult by the limited number of degrees of freedom for displacing the tip of the separation device relative to the target. A similar arrangement with a sub-atmospheric deposition chamber is described in Karger et al. (US patents 6,674,070 and 6825,463). Accordingly better coupling of separating/drop deposition device and targets are needed.

6.4 Materials and Methods

The housing for the SepDep II device was machined out of a 5 cm (diameter) x 8 cm aluminum cylinder to accept a co-axial piston fixed to an LC column and matrix delivery lines (Figure 6.1, Figure 6.2). As only one moving part is used, the probe of the chromatographic device (Vacuu-probe) is simple and rugged. Although manual operation can be envisioned using a 'pipette-like' interface, automation was provided by a New England Affiliated Technologies (NEAT/Danaher Motion, Salem, NH) 300 series multi-axis programmable motion controller interfaced to 3 linear stages for positioning and travel. Fused silica (360 μm o.d., 100 μm i.d., Polymicrotechnologies, Phoenix, AZ) was used for fluid transfer lines. Vacuum control was provided by a GraLab™ timer controller (Dimcogray, Centerville, Ohio). The device was evaluated using tryptic digest mixtures of phosphorylated β -casein using similar conditions previously described (Part 2.3.3). The in-lab proof-of-concept has

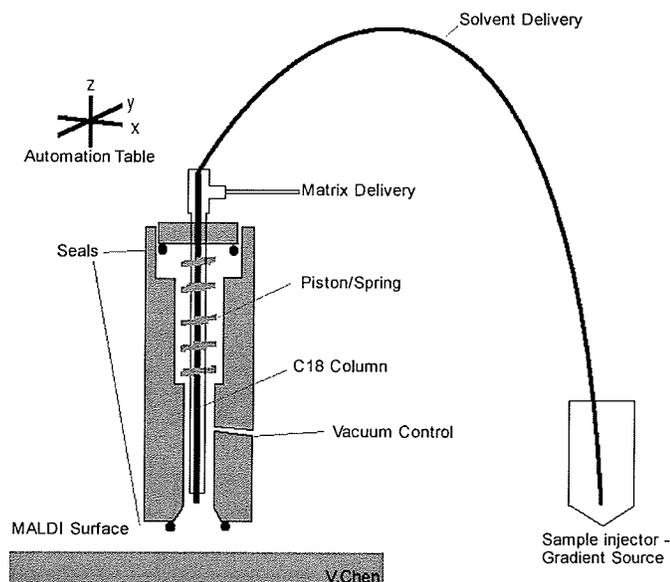


Figure 6.1: Schematic representation of the SepDep II Vacuu-flow sample preparation module for off-line LC-MALDI-MS. Vacuum control is provided by an electronic solenoid valve which monitors SepDep II. The MALDI target is used to enclose the miniaturized vacuum chamber when in contact with the Vacuu-probe. Once negative atmosphere is established, a vacuum activated piston is compressed and solvent flow through a LC column is generated with eluent deposited onto the MALDI surface. Matrix solution is mixed post-column using a micro-T junction.

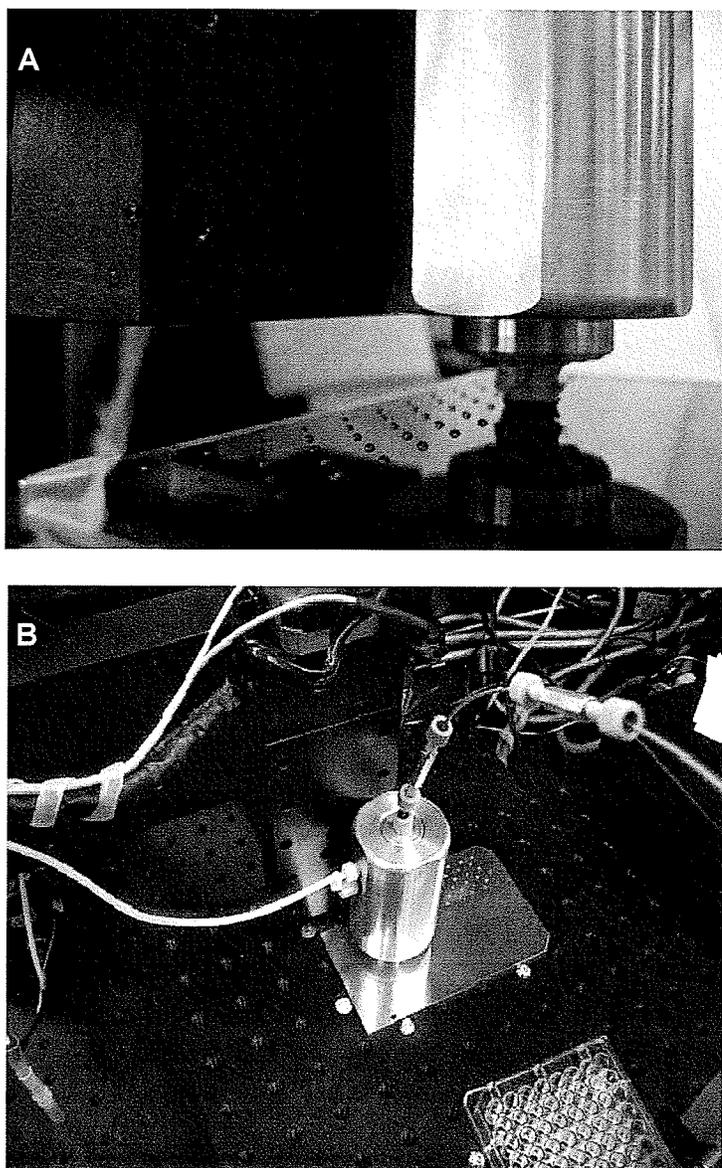


Figure 6.2: Images of the SepDep II prototype. (A) Deposition of reverse phase (C18) separated LC eluent, collected in 10 second intervals with co-deposition of DHB matrix onto a Bruker Anchorchip disposable target. (B) Top view of the probe interfaced to a stainless steel MALDI surface.

been demonstrated using C18 chromatography using phosphorylated protein standards. Investigations to date have demonstrated the device is capable of arraying 300-420 sample fractions unattended in 40 minutes at significantly higher throughput than the original SepDep device. A hydrophobic MALDI surface such as the Bruker Disposable Anchorchip™ MALDI Targets reduced the size of the sample spot and negated the need for the hydrophobic PapPen barrier (described in Part 2.3.6). Physical contact of the Vacuu-probe did not interfere with off-line sample collection at flow rates up to 60 $\mu\text{L}/\text{min}$. Analysis of phosphorylated β -casein led to the identification of all predicted PTMs; including tetra-phosphorylated peptide ELEELNVPGEIVEpSLpSpSpSEESITR (A.A. 17-40, Figure 6.3).

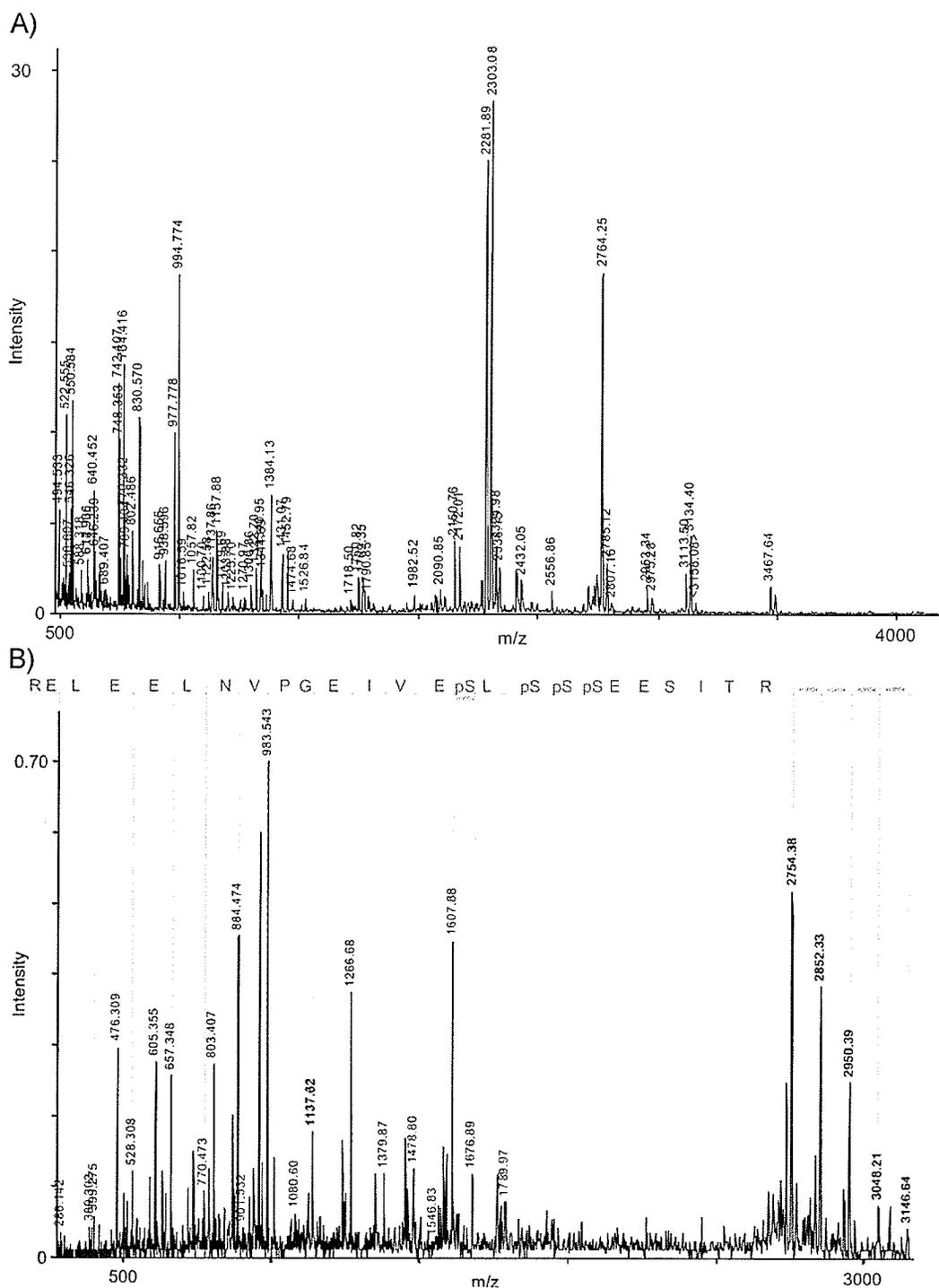


Figure 6.3: Identification of tetra-phosphorylated peptide. A) MS of tryptic β -casein digest mixture. B) MS/MS β -casein peptide (m/z 3146.64) obtained from SepDep II fraction O11 (fraction number 347). Fragmentation pattern of peptide ELEELNVPGEIVEpSLpSpSpSEESITR (A.A. 17-40). B-series ions labeled, along with consecutive neutral losses of -98 u (at m/z 3048, 2950, 2852, 2754) are noted.

6.5 SUMMARY OF THE INVENTION

[3] In one embodiment of the invention there is provided a sample deposition device comprising a housing which in turn comprises a vacuum chamber, a sealable opening communicating with the vacuum chamber, a vacuum inlet, a sample inlet and a sample outlet, the sample outlet located in the vacuum chamber, and wherein the sealable opening is sealed by a target surface when the housing is placed in sample deposition position on the target surface. A sample is drawn through the sample outlet and deposited on the target surface when a vacuum is applied in the vacuum chamber.

[4] In a further embodiment of the invention a separating chamber is coupled to the sample inlet and the sample outlet thereby providing a means for separating/purifying samples comprising different types of molecules such as a protein sample for example.

[5] In another aspect of the invention there is provided a method for depositing drops of a sample on a target surface said method comprising: providing a target surface, providing a vacuum activated sample depositing device wherein the target surface is a sealing member of the device, applying a vacuum to the device to draw a sample towards the target surface and effect deposition of one or more drops and releasing the vacuum.

[6] By using the target surface as a sealing member, the vacuum activated drop deposition device of the present invention provides a simplified set up for separating and depositing sample drops in a rapid and flexible manner.

6.5.1 BRIEF DESCRIPTION OF THE DRAWINGS

[7] Further features and advantages of the present invention will become apparent from the following detailed description, taken in combination with the appended drawings, in which:

[8] Figure 6.4 is a cross-sectional view of an embodiment of the device of the present invention;

[9] Figure 6.5A shows an embodiment of the invention in which capillary electrophoresis is used for the separation of molecules;

[10] Figure 6.5B shows an embodiment of the invention in which capillary electrophoresis is used for the separation of molecules wherein the target surface is part of the electric circuitry;

[0011] Figure 6.6 is cross-sectional view of an embodiment of the device shown with a removable lid;

[12] Figure 6.7 is a cross-sectional view of an embodiment of the device showing housings of different dimensions occupying 15 increasingly larger surface areas on the target surface;

[13] Figure 6.8 is a perspective view of an embodiment of the device in which a second vacuum chamber is shown;

[14] Figure 6.9A is an embodiment of the invention using a ball valve; and

[15] Figure 6.9B is an embodiment of the invention using a piston.

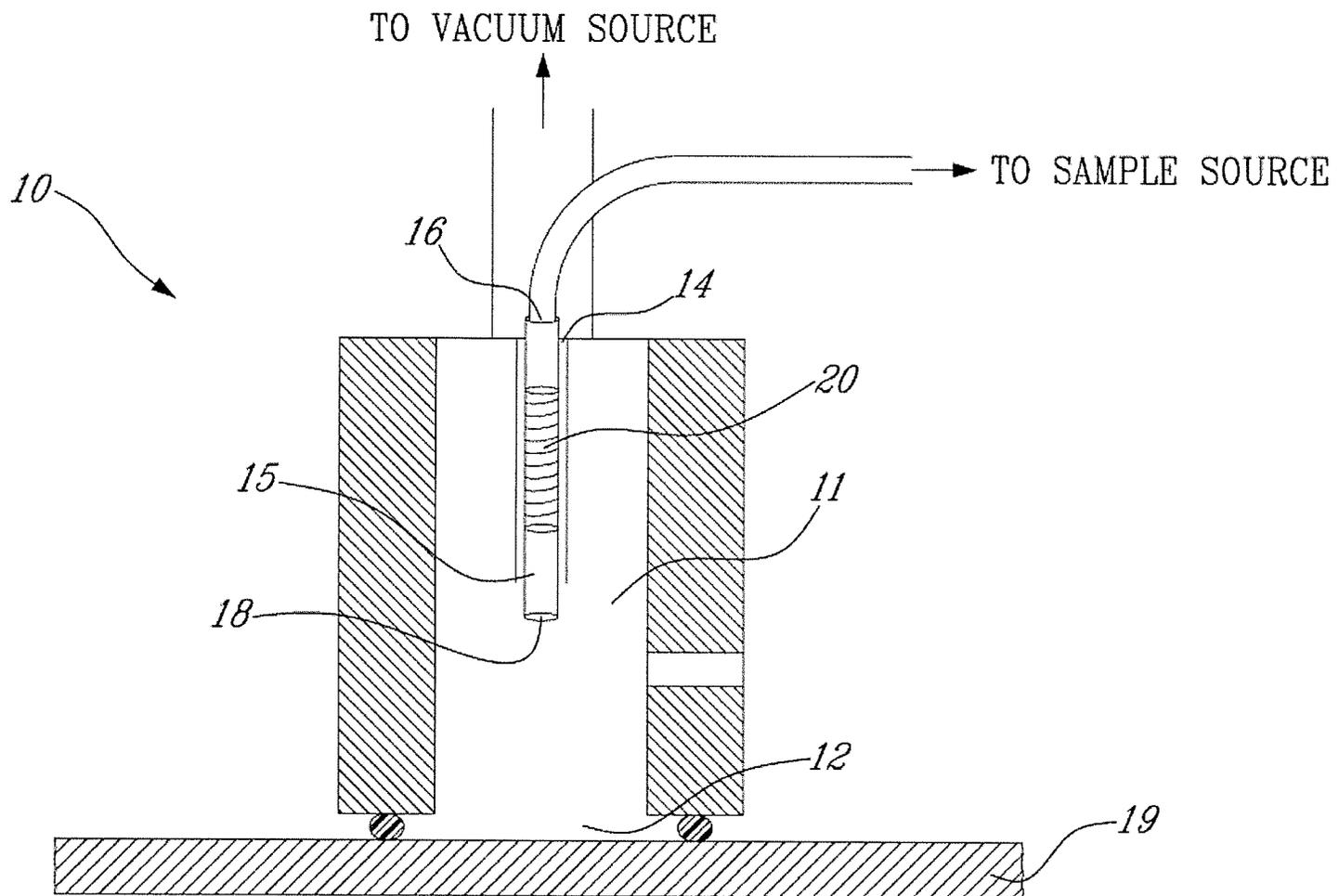


Figure 6.4

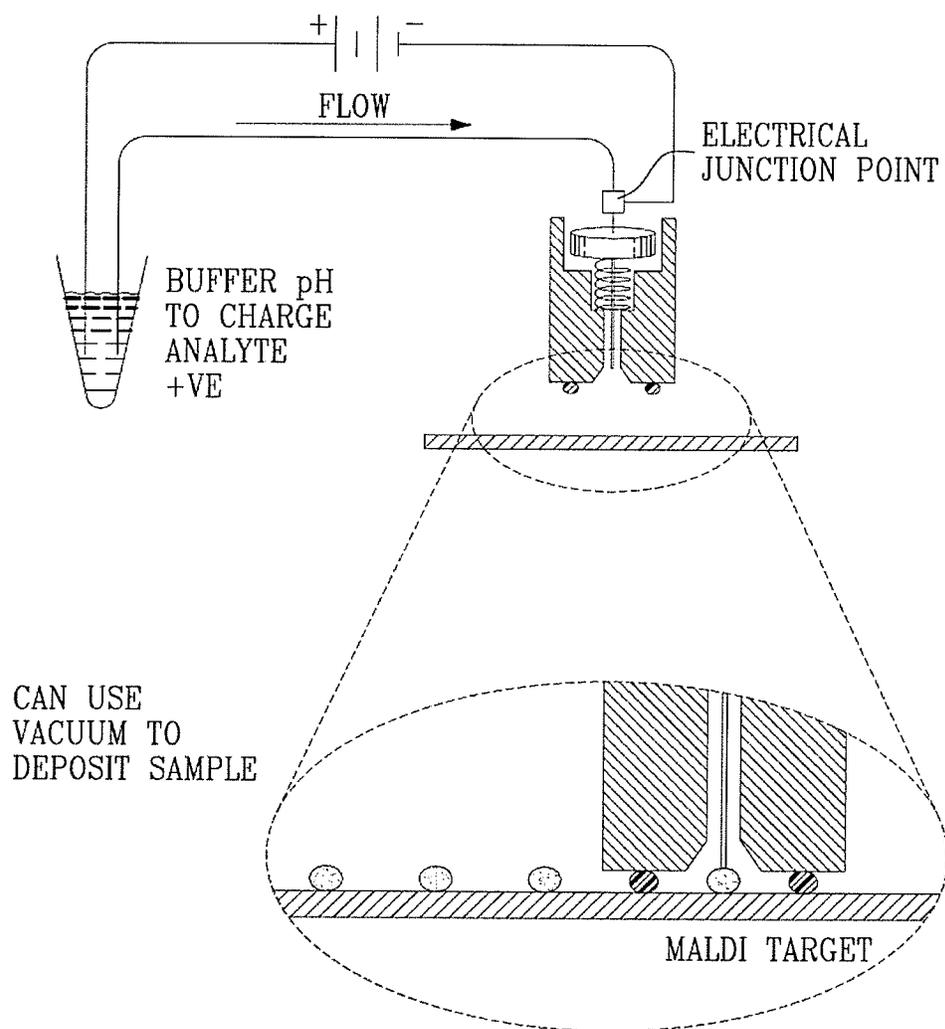


Figure 6.5A

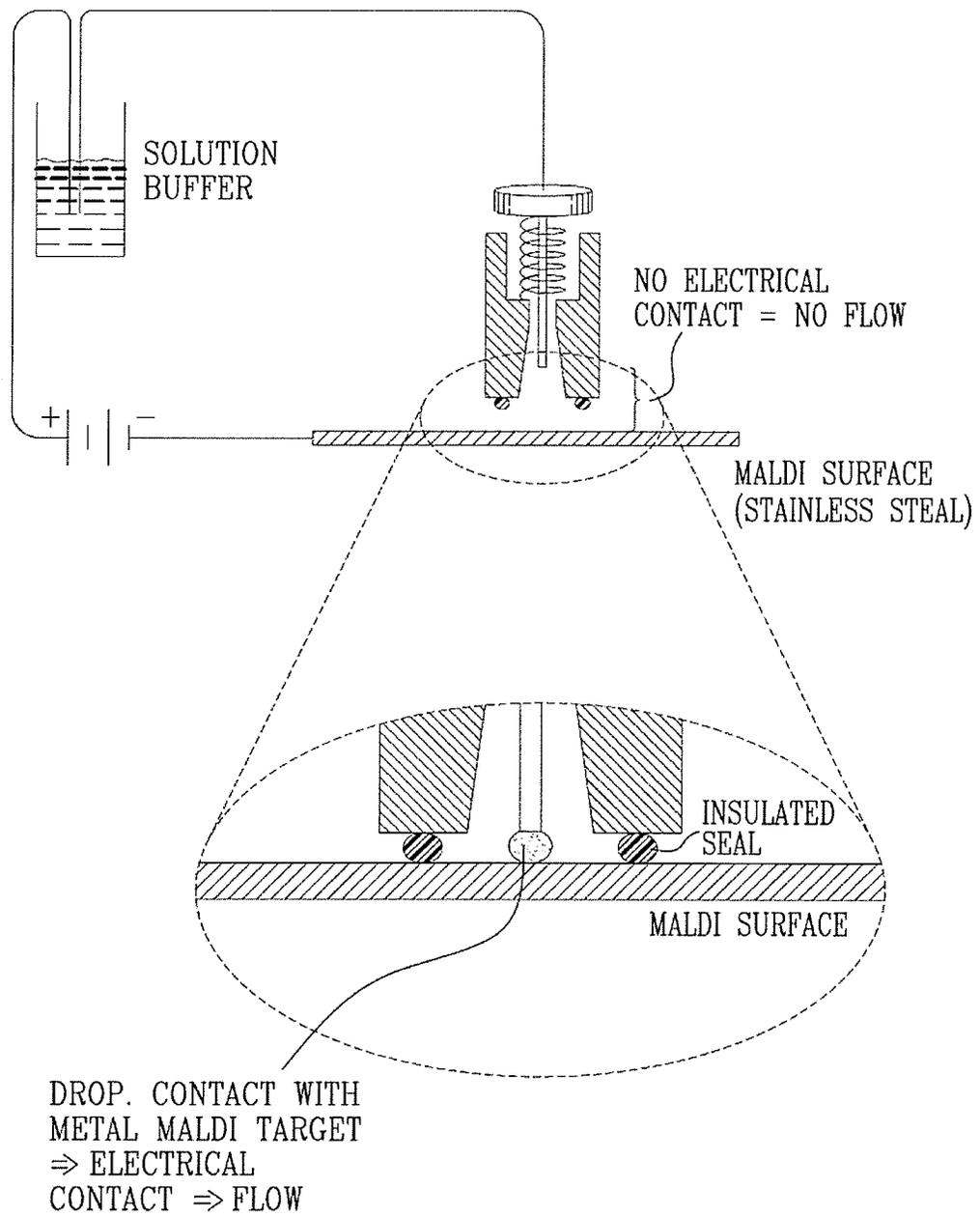
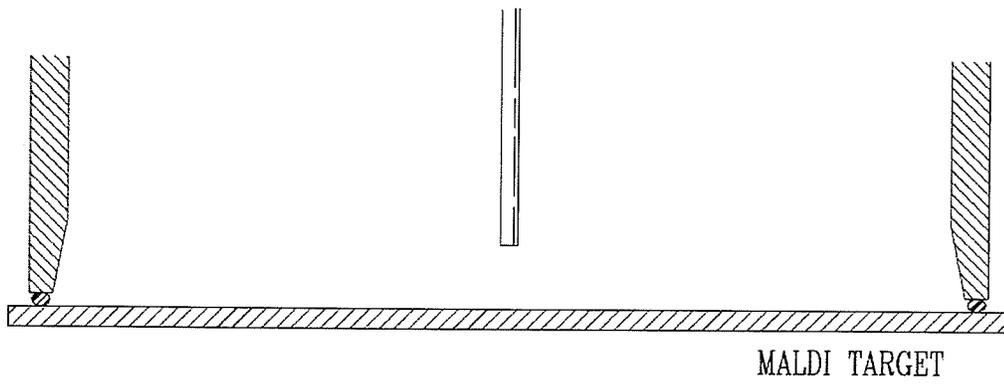
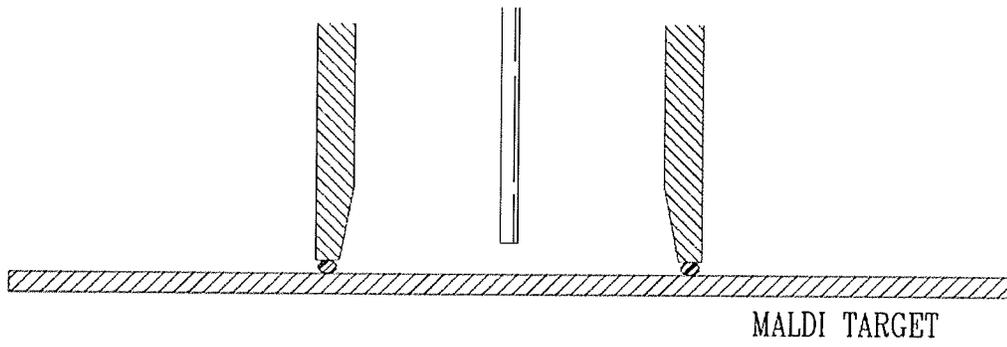
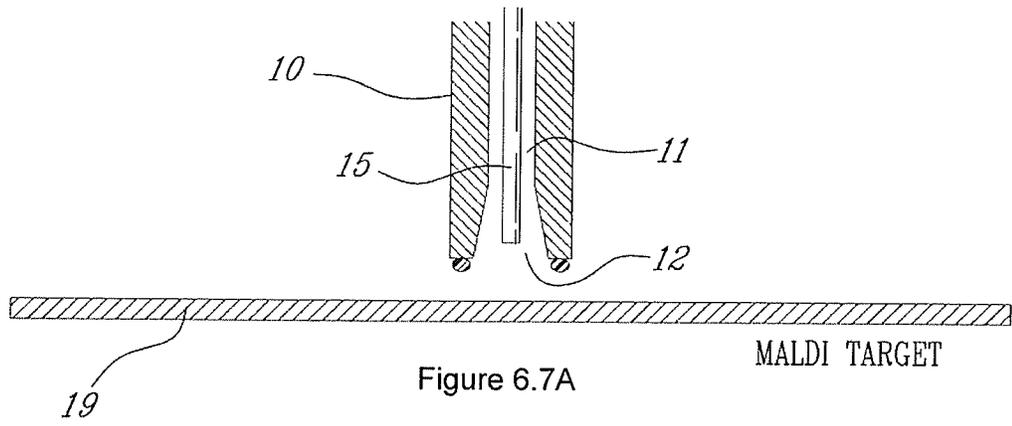


Figure 6.5B



6.5.2 DETAILED DESCRIPTION OF THE INVENTION

[16] A vacuum activated sample depositing/separating device is provided for deposition of samples on a target surface such as a 25 MALDI target surface for mass spectral analysis of the sample.

[17] An embodiment of the sample depositing device of the present invention is shown in Figure 6.4. The device comprises a housing 10 which in turn comprises a vacuum chamber 11, a sealable opening 12 communicating with the vacuum chamber, a sample channel 14 comprising a sample inlet 16 connected to a sample source and a sample outlet 18 located inside the vacuum chamber.

[18] The sample can be circulated directly in the sample channel 14 but, alternatively, it is also possible to insert a sample duct 15 within the sample channel or directly within the vacuum chamber. The sample duct may be adapted to help control the flow of the sample depending for example on the viscosity of the sample. Thus the size, shape and material of the duct may be chosen as function of the sample to be deposited. Furthermore the duct can be removed to be cleaned or replaced. When a sample duct is used, the sample channel may serve as a vacuum inlet which connects the vacuum chamber to a vacuum source. However, it will be appreciated that the vacuum inlet can be located anywhere along the vacuum chamber.

[19] In operation, that is, when a drop of sample is deposited, 15 the housing is placed on the target surface 19 at a desired location such that the target surface serves as a sealing member for the sealable opening 12. When a vacuum is applied within the vacuum chamber, the sample is drawn through the sample outlet and deposited on the surface. The use of the target surface as a sealing 20 member for the vacuum chamber advantageously allows the housing to be easily displaced over the target surface.

[20] In a preferred embodiment a separation chamber 20 is positioned between and connected with sample inlet 16 and sample outlet 18 such that when a vacuum is applied the sample is drawn into the separation chamber through sample inlet 16 and ultimately exiting the separation chamber through sample outlet 18 to be deposited on target surface 19. The separation chamber may be part of sample duct 15. The separation chamber 20 may be made in whole or in part of any suitable material including but not limited to bendable material that may facilitate the positioning of the separating chamber within the housing.

[21] In one embodiment the sample comprises a mixture of protein of interest that is preferably separated prior to being deposited on a target surface in preparation for mass spectrometry analysis. The separation of the molecules can be achieved by various means such as chromatography and electrophoresis. Non-limiting examples include capillary electrophoresis,

hydrophobic chromatography, ion exchange chromatography, affinity chromatography, size exclusion chromatography, chromatofocusing, capillary electrochromatography, solid phase extraction and the like. Monolithic columns, which are made of a porous network scaffold, can advantageously be used since they exhibit lower packing density when compared with traditional chromatographic material, which uses microbeads, and therefore allow the use of lower vacuum (closer to atmospheric pressure). Thus the separation chamber 20 may consist of a chromatographic column or capillary.

[22] When using capillary electrophoresis the separation may be achieved by applying a voltage at the sample source and the separation chamber (see Figure 6.5A). Once the analytes (macromolecules) are separated into "zones" the deposition of the drops may be effected by applying the vacuum. However in one particularly advantageous realization of the invention, the separation voltage may be applied between the sample and the target surface thereby producing a flow and separation within the separation chamber (see Figure 6.5B). The current is established when the drop makes contact with the target surface. In this last embodiment a vacuum may or may not be applied.

[0023] In one embodiment, the vacuum chamber can be sealed by introducing separation chamber 20 directly into the vacuum chamber or through sample channel 14 and with a sealing member such as a rubber "O" ring. In another embodiment, and referring to Figure 6.6, a removable lid 22

comprising sample channel 14 can be used. A biasing member 24, such as a spring, when no vacuum is being applied, can maintain the lid in an open position. Sealing members are preferably coupled to the lid to provide a tight seal between the lid and the housing.

[24] The separating chamber can be removeably attached to the lid or allowed to slide through the sample inlet. In the first case the position of the separating chamber relative to the target surface, and therefore of sample outlet 18 is controlled by displacement of the lid. For example, if a biasing member is maintaining the lid open, a force can be applied on the lid which will cause the vacuum chamber to be sealed and bring the sample outlet in proximity to the target surface for sample deposition. In the second case the position of the separating chamber can be controlled manually by a user or automatically by using a motorized robot for example by sliding the separating chamber in the sample inlet to a desired position inside the vacuum chamber.

[25] While the connection to the vacuum source is shown in Figures 6.4 and 6.6 as being made through sample channel 14, it will be appreciated that it can be located at any other place along the vacuum chamber. It will also be appreciated that in the latter embodiment, a proper sealing member should be provided between the separation chamber and/or sample duct and sample chamber 14. For example, in the case where the separation chamber is a

chromatographic column, the column could be inserted in a rubber "O" ring. Other types of sealing members are well known in the art.

[26] The deposition of a sample drop may be effected as follows: the housing is brought in juxtaposition of the target surface by placing a sealable opening 12 on the target surface, at a position where a drop is to be deposited, thereby effecting the sealing of this end of the vacuum chamber. The vacuum is then applied to draw the sample through the sample inlet and sample outlet and effect deposition of a sample drop. The cycle is completed by releasing the vacuum and displacing the device to the next drop deposition position. While the release of the vacuum can be affected by pulling the housing away from the target surface, in a preferred embodiment a valve or opening connected to the vacuum chamber is provided for controlled release of the vacuum so as to avoid disturbing the deposited drop. If a separating chamber is being used, the application of a vacuum creates a flow of the sample through the separation chamber thereby effecting separation of the molecules within the sample.

[27] The vacuum intensity and its duration can be adjusted such as to effect the deposition of a desired volume of sample and to control the speed at which the sample is carried through the separation chamber. The volume deposited at each spot on the surface may also be controlled by changing the distance between sample outlet 18 and the target surface.

[28] The surface occupied by the housing at the point of contact with the target surface can be tailored to the need of a particular application. In one aspect, the contact surface area may be dictated by the desired spacing between drops. However, the surface area occupied by the housing on the surface is only limited by the surface area of the target surface (see Figure 6.7). In one embodiment, the surface area occupied by the housing on the target surface is sufficiently large so that the volume of the vacuum chamber may allow the displacement of the separation chamber to apply more than one drop at different positions while the housing remains stationary on the target surface.

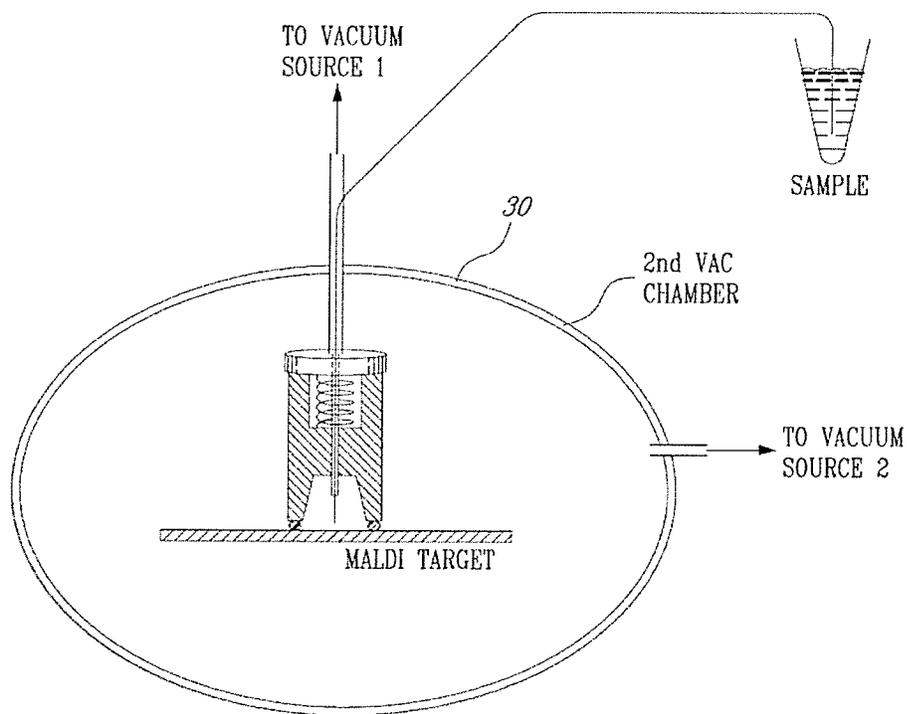


Figure 6.8

[29] It is desirable to release the vacuum after drop(s) deposition in a controlled manner so as to avoid disturbing the drop(s). This can be achieved by controlling the release of the vacuum with valves. In an embodiment of the invention this can also be achieved by providing a second vacuum chamber 30 (Figure 6.8) in which the target surface and the housing can be placed. The vacuum in the second vacuum chamber is maintained at a pressure slightly higher than the pressure within the vacuum chamber of the device but lower than the atmospheric pressure. This arrangement allows the release of the pressure within the vacuum chamber of the device with minimal disturbance of the drop. The pressure in the second vacuum chamber can be adjusted to produce a sample flow albeit at a lower rate than the flow generated by the vacuum in the vacuum chamber 11 or 10 alternatively it can be adjusted to be insufficient to generate solvent flow.

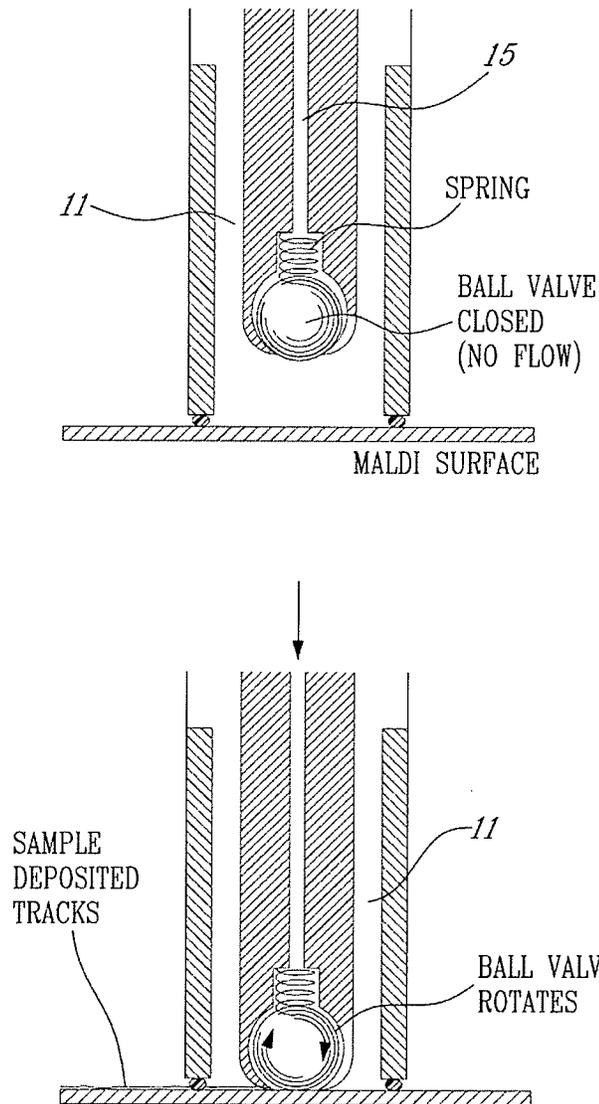


Figure 6.9A

[0030] Other embodiments of the invention are depicted in Figure 6.9A and 6.9B. In Figure 6.9A the sample channel 15 within vacuum chamber 11 has a ball valve to control the delivery of the sample on the target surface. When the device is in the delivery position i.e. against the target surface the ball is displaced upwards and the sample can flow onto the surface. When the device is not in the delivery position a biasing means such as a spring,

pushes the ball against the opening and prevents the flow of the sample. This embodiment advantageously allows the sample to be spread on the surface or to be deposited as discrete spots.

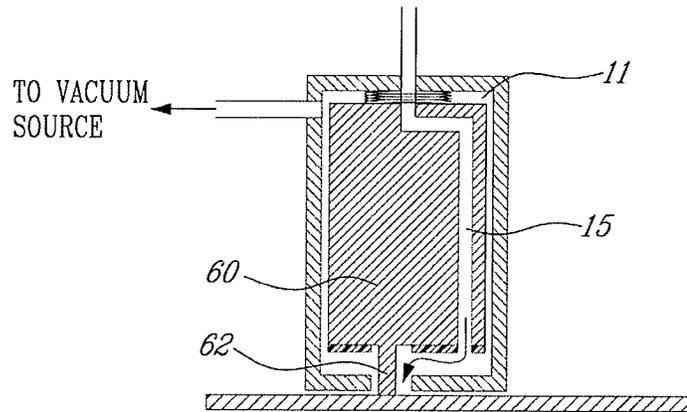
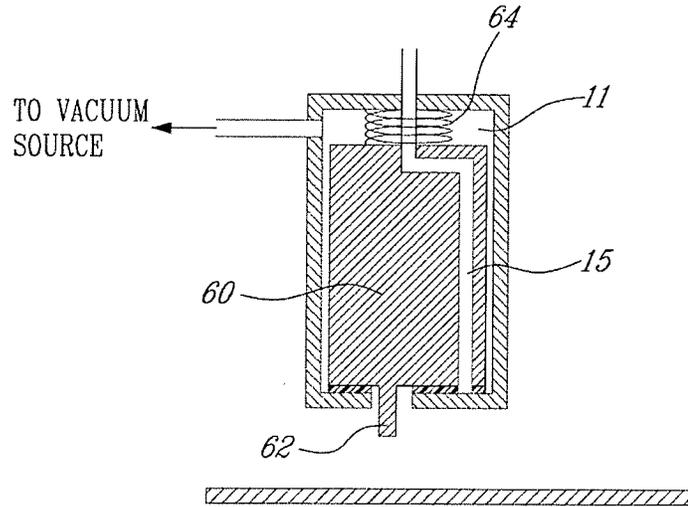


Figure 6.9B

[0031] Figure 6.9B exemplifies yet another embodiment wherein the actuation of the sample flow is effected by a piston 60 having a piston contact bar 62. When the device is contacting the target surface the piston is displaced

upwards and the bottom of the housing comes in contact with the target surface. The upward displacement of the piston enables the vacuum source and creates a sample flow for deposition on the surface. When the device is not in a delivery position the piston is biased against the bottom of the housing by biasing means 64 prevents the vacuum from creating a sample flow. It will be appreciated that when the device is in a second vacuum chamber, the bottom of the housing need not come in contact with the target surface since the vacuum is external to the housing.

[0032] While the device can easily be hand held and manipulated by a user, it will be appreciated that the various steps in the deposition of a sample drop can be performed mechanically by a computer-controlled robot. In particular, mechanical manipulation is advantageous in the embodiment where a second vacuum chamber is used and whereby manipulation of the device is performed inside the second vacuum chamber by mechanical means. It will further be appreciated that not only the displacement of the housing over the target surface may be automated but any aspect of the apparatus that is amenable to mechanical control. For example, displacement of the separating chamber within the vacuum chamber, particularly in the embodiment in which the vacuum chamber is large enough to allow deposition of several drops at different positions on the target surface, control of vacuum release, vertical adjustment of the sample outlet and the like. Furthermore, positioning of the housing, and therefore the drops, on the target surface may also be

accomplished by displacing the surface while maintaining the housing stationary. Displacement of both the housing and the target surface is also possible.

[0033] In one embodiment of the invention the deposition device can be coupled to an automated liquid handler (liquid handling robot) such as to provide motion control of the device as well as regulation of the solvent/sample liquid flow into a sample duct. It will be appreciated that in the case where a liquid handler is used, the liquid flow is controlled by the negative pressure within the device and the liquid pressure exerted by the pump of the liquid handler.

[0034] In a preferred embodiment the deposition device of the present invention is used to separate and deposit macromolecules on the surface of a MALDI target. By macromolecules it is meant proteins, peptides, oligonucleotides and the likes. MALDI target plates are well known in the art and may consist but are not limited to a stainless steel plate such as the Bruker Daltonics 384 target.

[0035] While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the

invention and including such departures from the present disclosures as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features herein before set forth, and as follows in the scope of the appended claims.

6.5.2 CLAIMS:

[1] A sample deposition device comprising: a housing comprising:

- i) a vacuum chamber;
- ii) a sealable opening communicating with said vacuum chamber;
- iii) a vacuum inlet;

iv) a sample inlet and a sample outlet, said sample outlet located in said vacuum chamber; and wherein said sealable opening is sealed by a target surface when said housing is placed in sample deposition position on said target surface and wherein a sample is drawn through said sample outlet and deposited on said target surface when a vacuum is applied in said vacuum chamber.

[2] The sample deposition device as claimed in claim 1 further comprising a sample duct adapted to fit in said sample channel.

[3] The sample deposition device as claimed in claim 1 or 2 further comprising a sample separating chamber positioned between said sample inlet and said

sample outlet such that said sample is drawn through said separation chamber when a vacuum is applied.

[4] The sample deposition device as claimed in claim 3 wherein said separating chamber is comprised within said sample duct.

[5] The sample deposition device as claimed in any one of claims 1-4 further comprising a removable lid wherein said lid is sealably connected to said housing when closed.

[6] The sample deposition device as claimed in claim 5 wherein said sample inlet is comprised in said lid.

[7] The sample deposition device as claimed in claim 5 or 6 further comprising a biasing member for biasing said lid in an open position.

[8] The sample deposition device as claimed in claim 7 wherein said biasing means is a spring.

[9] The sample deposition device as claimed in any one of claims 5-8 wherein said separating chamber is removably attached to said lid whereby said sample outlet is brought in sample deposition position when said lid is closed.

[10] The sample deposition device as claimed in any one of claims 1-9 wherein said separating chamber is a chromatographic column.

[11] The sample deposition device as claimed in claim 10 wherein said chromatographic column comprises said sample inlet and outlet.

[12] The sample deposition device as claimed in claim 11 wherein said column can be displaced laterally within said vacuum chamber such as to enable deposition of two or more sample drops at two or more positions on said target surface while said housing is maintained stationary.

[13] The sample deposition device as claimed in any one of claims 1-12 wherein said target surface is a MALDI target.

[14] The sample deposition device as claimed in any one of claims 1-13 further comprising a vacuum outlet for controllably releasing said vacuum.

[15] The sample deposition device as claimed in any one of claims 1-14 wherein said target plate and said vacuum outlet are comprised within a second vacuum chamber having an adjustable internal pressure below atmospheric pressure and above a pressure in said first vacuum chamber.

[16] The sample deposition device as claimed in any one of claims 1-15 wherein said sample comprises proteins of interest.

[17] A sample deposition system comprising the sample deposition device as claimed in any one of claims 1-16 and an automated liquid handler wherein said automated liquid handler is operationally coupled to said sample deposition device.

[18] A method for depositing drops of a sample on a target surface said method comprising:

- a) providing a target surface
- b) providing a vacuum activated sample depositing device wherein said target surface is a sealing member of said device;
- c) applying a vacuum to said device to draw said sample towards said target surface and effect deposition of one or more drop; and
- d) releasing said vacuum.

[19] The method as claimed in claim 18 further comprising a step of separating components of said sample while said sample is being drawn by said vacuum.

[20] The method as claimed in claim 18 or 19 wherein said sample comprises proteins of interest.

[21] The method as claimed in any one of claims 18-20 wherein said target surface is a MALDI target.

[22] The method as claimed in any one of claims 18-21 wherein said vacuum activated sample depositing device is as claimed in any one of claims [1-16, 20].

[23] The method as claimed in claim 22 wherein said sealable opening occupies a part of said target surface when said housing is in a sample deposition position.

[24] The method as claimed in claim 23 wherein said sealable opening occupies substantially all of said target surface when said housing is in a sample deposition position and wherein said sample outlet is displaced within the vacuum chamber to deposit sample at two or more positions on said target surface.

7.0 Supplementary Information:
The targeted identification of phosphorylated peptides by Off-line HPLC-
MALDI-MS/MS using LC Retention Time Prediction

Protein identified within S2 cell anti-phosphotyrosine immunoprecipitates by
off-line HPLC data-dependent MALDI-TOF-TOF-MS/MS.

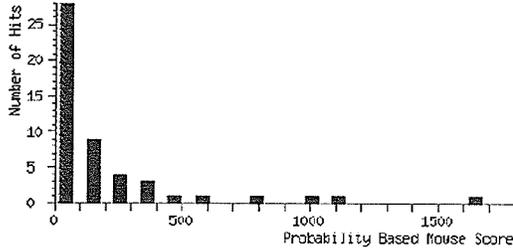
(MATRIX)
(SCIENCE) **Mascot Search Results**

Significant hits: gi 111042	hrp48.1 [Drosophila melanogaster]
gi 19921124	SCAR CG4636-PA [Drosophila melanogaster]
gi 24650831	Heterogeneous nuclear ribonucleoprotein at 98DE CG9983-PA, isoform A [Drosophila
gi 21355219	SH3PX1 CG6757-PA [Drosophila melanogaster]
gi 17137386	Abelson Interacting Protein CG9749-PA [Drosophila melanogaster]
gi 18094	Hrb87F [Drosophila melanogaster]
gi 14161109	PDGF/VEGF receptor [Drosophila melanogaster]
gi 13430414	actin E2 [Drosophila virilis]
gi 21357451	specifically Rac1-associated protein 1 CG4931-PA [Drosophila melanogaster]
gi 158739	beta-1 tubulin
gi 17426899	Eps-15 protein [Drosophila melanogaster]
gi 1495594	poly(A)-binding protein [Drosophila melanogaster]
gi 3869204	Cortactin [Drosophila melanogaster]
gi 24646107	CG6946-PB, isoform B [Drosophila melanogaster]
gi 24653266	Fak-like tyrosine kinase CG3969-PA, isoform A [Drosophila melanogaster]
gi 4377461	Dos protein [Drosophila melanogaster]
gi 39752635	LF10071p [Drosophila melanogaster]
gi 21428576	LD43495p [Drosophila melanogaster]
gi 17647889	Ribosomal protein S17 CG3922-PB [Drosophila melanogaster]
gi 1498137	TamA [Drosophila melanogaster]
gi 157658	heat shock protein cognate 72
gi 8072217	Dscam [Drosophila melanogaster]
gi 111038	hrp40.1 [Drosophila melanogaster]
gi 3255957	EG:132E8.1 [Drosophila melanogaster]
gi 3834641	Drongo [Drosophila melanogaster]
gi 24663131	ypsilon schachtel CG5654-PA [Drosophila melanogaster]
gi 17647885	Ribosomal protein L18A CG6510-PA [Drosophila melanogaster]
gi 18476	unnamed protein product [Drosophila melanogaster]
gi 18858095	Ribosomal protein S28b CG2998-PA [Drosophila melanogaster]
gi 24645526	CG8516-PA [Drosophila melanogaster]
gi 24648551	Ribosomal protein S30 CG15697-PB, isoform B [Drosophila melanogaster]
gi 24658098	quaking related 58E-2 CG5821-PA [Drosophila melanogaster]
gi 16690788	Short stop/Kakapo long isoform [Drosophila melanogaster]
gi 38047863	similar to Drosophila melanogaster CG3195 [Drosophila yakuba]
gi 3337433	transitional endoplasmic reticulum ATPase TER94 [Drosophila melanogaster]
gi 24646982	CG6904-PC, isoform C [Drosophila melanogaster]
gi 20129705	Ribosomal protein L21 CG12775-PA [Drosophila melanogaster]
gi 12958781	no-mechanoreceptor potential A long isoform [Drosophila melanogaster]
gi 24640206	Ribosomal protein L17 CG3283-PC, isoform C [Drosophila melanogaster]
gi 1755821	ena polypeptide
gi 121356037	CG9780-PA, isoform A [Drosophila melanogaster]
gi 115010476	GH06265p [Drosophila melanogaster]
gi 24653903	fusilli CG8205-PA, isoform A [Drosophila melanogaster]
gi 17647897	Ribosomal protein S5a CG8922-PA [Drosophila melanogaster]
gi 18860517	Mec2 CG7635-PA [Drosophila melanogaster]
gi 116197837	GH10594p [Drosophila melanogaster]
gi 1228974	vinculin [Drosophila melanogaster]
gi 8250181	D-Titin [Drosophila melanogaster]
gi 17576913	E protein transcript B [Drosophila melanogaster]
gi 3056723	translation initiation factor eIF4G [Drosophila melanogaster]
gi 18444	unnamed protein product [Drosophila melanogaster]
gi 13337431	projectin [Drosophila melanogaster]
gi 117985987	belle CG9748-PA [Drosophila melanogaster]
gi 119920688	CG15436-PA [Drosophila melanogaster]
gi 456984	ryanodine receptor, calcium release channel [Drosophila melanogaster, Peptide, 5
gi 24581733	CG15627-PA [Drosophila melanogaster]
gi 4914494	pacman protein [Drosophila melanogaster]
gi 18857967	CG10777-PB [Drosophila melanogaster]
gi 117998549	no mechanoreceptor potential C CG11020-PA, isoform A [Drosophila melanogaster]
gi 18488	unnamed protein product [Drosophila melanogaster]
gi 38048361	similar to Drosophila melanogaster CG8495 [Drosophila yakuba]
gi 117945862	RE39037p [Drosophila melanogaster]
gi 16225885	defective chorion-1 fc177 protein precursor [Drosophila yakuba]
gi 4056674	plexin A [Drosophila melanogaster]
gi 24645017	CG9626-PA [Drosophila melanogaster]
gi 38048391	similar to Drosophila melanogaster RpS18 [Drosophila yakuba]
gi 24655594	Dynein heavy chain at 62B CG15804-PA, isoform A [Drosophila melanogaster]
gi 62484380	lingerer CG8715-PB, isoform B [Drosophila melanogaster]
gi 349669	cytoplasmic dynein heavy chain
gi 433182	receptor protein tyrosine phosphatase
gi 24640934	CG2989-PA [Drosophila melanogaster]
gi 24580801	CG5080-PA, isoform A [Drosophila melanogaster]
gi 6644386	dynein heavy chain [Drosophila melanogaster]
gi 17864174	maverick CG1901-PA, isoform A [Drosophila melanogaster]
gi 13492045	SPRINT-a [Drosophila melanogaster]
gi 7682380	gp150 [Drosophila virilis]
gi 24584386	crinkled CG7595-PA, isoform A [Drosophila melanogaster]
gi 20151585	LD24077p [Drosophila melanogaster]

[gi|6822326](#) VAV protein [Drosophila melanogaster]
[gi|24642261](#) CG12379-PA [Drosophila melanogaster]
[gi|13892027](#) leucine rich repeat protein GP150 [Drosophila melanogaster]
[gi|24654746](#) CG1231-PA [Drosophila melanogaster]
[gi|2645435](#) CHD3 [Drosophila melanogaster]
[gi|5679048](#) GM13640p [Drosophila melanogaster]
[gi|24652047](#) hikaru genki CG2040-PB, isoform B [Drosophila melanogaster]
[gi|2773363](#) microtubule binding protein D-CLIP-190 [Drosophila melanogaster]

Probability Based Mowse Score

Score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event. Individual ions scores > 32 indicate identity or extensive homology ($p < 0.05$).



Peptide Summary Report

[Switch to Protein Summary Report](#)

To create a bookmark for this report, right click this link: [Peptide Summary Report \(SampleSetID: 42, AnalysisID: 82, Path=Chi-ChIP\0723\)](#)

Error tolerant

1. [gi|11042](#) Mass: 41004 Total score: 1648 Peptides matched: 23
hrp48.1 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> 29	645.23	644.23	644.27	-0.05	0	25	1	YFCR + Carbamidomethyl (C)
<input checked="" type="checkbox"/> 84	684.31	683.31	683.34	-0.03	0	22	1	TFNR
<input checked="" type="checkbox"/> 196	1032.47	1031.46	1031.52	-0.06	1	35	1	TFNRYGK
<input checked="" type="checkbox"/> 243	1197.53	1196.52	1196.60	-0.07	0	58	1	TIDPKPCNPR + Carbamidomethyl (C)
<input checked="" type="checkbox"/> 244	1197.57	1196.56	1196.60	-0.04	0	(14)	3	TIDPKPCNPR + Carbamidomethyl (C)
<input checked="" type="checkbox"/> 243	1206.49	1205.48	1205.54	-0.06	0	130	1	QEGASNYGAGPR
<input checked="" type="checkbox"/> 249	1206.52	1205.51	1205.54	-0.03	0	(13)	3	QEGASNYGAGPR
<input checked="" type="checkbox"/> 250	1206.53	1205.53	1205.54	-0.02	0	(23)	1	QEGASNYGAGPR
<input checked="" type="checkbox"/> 291	1453.63	1452.63	1452.72	-0.09	0	61	1	VTEVVMYDQEK
<input checked="" type="checkbox"/> 357	1717.83	1716.82	1716.90	-0.08	0	107	1	VFLGGLPSNVTEIDLR
<input checked="" type="checkbox"/> 364	1769.69	1768.68	1768.76	-0.08	0	170	1	DGSGGCNSNNSTVGGAYGK
<input checked="" type="checkbox"/> 391	1898.82	1897.82	1897.91	-0.09	0	166	1	AQAWATGGPSTTGPVGGMPR
<input checked="" type="checkbox"/> 392	1898.87	1897.86	1897.91	-0.05	0	(9)	1	AQAWATGGPSTTGPVGGMPR
<input checked="" type="checkbox"/> 393	1898.89	1897.88	1897.91	-0.03	0	(51)	1	AQAWATGGPSTTGPVGGMPR
<input checked="" type="checkbox"/> 399	1936.89	1935.89	1935.97	-0.08	0	126	1	LFVGGLSWETTOENLSR
<input checked="" type="checkbox"/> 403	1953.83	1952.82	1952.91	-0.09	1	72	1	FGDIIDCVVMKNESGR + Carbamidomethyl (C)
<input checked="" type="checkbox"/> 434	2122.02	2121.01	2121.09	-0.07	1	50	1	GKLFVGGLSWETTOENLSR
<input checked="" type="checkbox"/> 456	2299.93	2298.93	2299.04	-0.11	0	92	1	GFGFLSFEESVSVEHVITNER
<input checked="" type="checkbox"/> 490	2608.92	2607.91	2608.02	-0.11	0	146	1	SGSEYDYGGSYDYDYSNYVK
<input checked="" type="checkbox"/> 511	2910.34	2909.33	2909.42	-0.09	0	109	1	GFGFVTFADPTNVNHLVQNGPHTLDGR
<input checked="" type="checkbox"/> 512	2910.36	2909.35	2909.42	-0.07	0	(56)	1	GFGFVTFADPTNVNHLVQNGPHTLDGR
<input checked="" type="checkbox"/> 523	3153.42	3152.41	3152.55	-0.14	1	16	1	SRGFGFVTFADPTNVNHLVQNGPHTLDGR
<input checked="" type="checkbox"/> 535	3796.40	3795.39	3795.55	-0.16	1	262	1	SGSEYDYGGSYDYDYSNYVKQEGASNYGAGPR

Proteins matching the same set of peptides:
[gi|66804007](#) Mass: 44742 Total score: 1648 Peptides matched: 23
 heterogeneous nuclear ribonucleoprotein [Drosophila melanogaster]

2. [gi|19921124](#) Mass: 67013 Total score: 1074 Peptides matched: 18

SCAR CG4636-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
157	992.50	991.50	991.57	-0.07	1	18	2	IRVPHNTR
<input checked="" type="checkbox"/> 193	1014.56	1013.56	1013.60	-0.04	1	26	1	AIRDGITLR
<input checked="" type="checkbox"/> 213	1109.48	1108.47	1108.55	-0.07	0	23	1	MLPPFHDFR
<input checked="" type="checkbox"/> 224	1139.53	1138.53	1138.58	-0.05	0	66	1	VEDQQIFSR
<input checked="" type="checkbox"/> 225	1139.54	1138.53	1138.58	-0.05	0	(59)	1	VEDQQIFSR
<input checked="" type="checkbox"/> 226	1139.55	1138.54	1138.58	-0.03	0	(13)	1	VEDQQIFSR
<input checked="" type="checkbox"/> 227	1139.56	1138.55	1138.58	-0.02	0	(18)	1	VEDQQIFSR
<input checked="" type="checkbox"/> 256	1237.58	1236.57	1236.64	-0.07	1	49	1	KMLPPFHDFR
<input checked="" type="checkbox"/> 261	1243.54	1242.53	1242.60	-0.07	0	101	1	HAEDVFGELAR
<input checked="" type="checkbox"/> 271	1310.69	1309.68	1309.74	-0.05	0	107	1	NAAPLDVASILAR
<input checked="" type="checkbox"/> 371	1797.75	1796.74	1796.82	-0.08	0	25	1	FYTDPNYFFELWR
<input checked="" type="checkbox"/> 373	1826.86	1825.86	1825.95	-0.10	0	112	1	ALVHGETLMPNNVIYR
<input checked="" type="checkbox"/> 413	2015.95	2014.94	2015.04	-0.10	0	154	1	VTQLDSTVEEVLTDITR
<input checked="" type="checkbox"/> 463	2400.09	2399.08	2399.20	-0.12	0	87	1	SLANGEMQQPGQQNGVPHIVAPK
<input checked="" type="checkbox"/> 474	2441.24	2440.24	2440.34	-0.11	1	45	1	LAIKVIQLDSTVEEVLTDITR
<input checked="" type="checkbox"/> 514	2994.45	2993.44	2993.52	-0.08	0	175	1	SVYQQDELQSVLELTVNNTLTNIIR
<input checked="" type="checkbox"/> 520	3110.41	3109.40	3109.53	-0.13	0	7	2	NKPRPSQPPAPFNSGGGTPASNANTFTR
<input checked="" type="checkbox"/> 536	3869.77	3868.77	3868.95	-0.18	0	80	1	MSPNAAPPPPPPPVVEEGMGSNGHQLTRPHQILPK

3. [gi|24650831](#) Mass: 39014 Total score: 1040 Peptides matched: 16

Heterogeneous nuclear ribonucleoprotein at 98DE CG9983-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> 12	631.27	630.26	630.31	-0.05	0	27	1	AHEFEK
<input checked="" type="checkbox"/> 173	949.37	948.37	948.42	-0.05	0	53	1	GGGNFNHNR
<input checked="" type="checkbox"/> 199	1053.53	1052.52	1052.57	-0.05	0	86	1	LFIGGLDYR
<input checked="" type="checkbox"/> 200	1065.52	1064.51	1064.58	-0.07	1	38	1	SRPEKIDGR
<input checked="" type="checkbox"/> 206	1073.42	1072.42	1072.46	-0.05	0	80	1	QNDQQGGGGGR
<input checked="" type="checkbox"/> 235	1169.46	1168.45	1168.53	-0.08	0	71	1	MQPYQGGGGFK
<input checked="" type="checkbox"/> 238	1181.62	1180.61	1180.66	-0.05	1	79	1	KLFIGGLDYR
<input checked="" type="checkbox"/> 253	1482.67	1481.66	1481.73	-0.07	1	101	1	ALPKQNDQQGGGGGR
<input checked="" type="checkbox"/> 319	1554.62	1553.61	1553.70	-0.09	1	(93)	1	QNDQQGGGGGRGGGGR
<input checked="" type="checkbox"/> 320	1554.65	1553.64	1553.70	-0.07	1	118	1	QNDQQGGGGGRGGGGR
<input checked="" type="checkbox"/> 342	1627.80	1626.79	1626.86	-0.07	1	44	1	WGNIVDVVVMKDFR
<input checked="" type="checkbox"/> 414	2017.85	2016.85	2016.93	-0.08	0	37	1	GFGFITYSBSSMIDEAQK
<input checked="" type="checkbox"/> 447	2237.83	2236.82	2236.93	-0.11	0	170	1	AGGNGQNYGNNQGNNGGNR
<input checked="" type="checkbox"/> 461	2331.05	2330.05	2330.15	-0.10	1	69	1	GFAFVEFDDYDFVDKVVLLQK
<input checked="" type="checkbox"/> 506	2851.03	2850.02	2850.14	-0.12	0	53	1	GNMGGNYGNQGGNWNNGNNGNNR
<input checked="" type="checkbox"/> 526	3306.22	3305.22	3305.36	-0.15	1	14	1	AGGNRGNMGGNYGNQGGNWNNGNNGNNR

Proteins matching the same set of peptides:

- [gi|24650833](#) Mass: 38527 Total score: 1040 Peptides matched: 16
Heterogeneous nuclear ribonucleoprotein at 98DE CG9983-PE, isoform E [Drosophila melanogaster]
- [gi|24650836](#) Mass: 39014 Total score: 1040 Peptides matched: 16
Heterogeneous nuclear ribonucleoprotein at 98DE CG9983-PC, isoform C [Drosophila melanogaster]
- [gi|24650840](#) Mass: 38527 Total score: 1040 Peptides matched: 16
Heterogeneous nuclear ribonucleoprotein at 98DE CG9983-PF, isoform F [Drosophila melanogaster]

4. [gi|21385219](#) Mass: 63127 Total score: 832 Peptides matched: 10

SH3PX1 CG6757-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> 263	1249.50	1248.49	1248.54	-0.05	0	78	1	SDVCEGWEGK
<input checked="" type="checkbox"/> 331	1590.64	1589.64	1589.72	-0.09	1	116	1	SDVCEGWEGKNAR
<input checked="" type="checkbox"/> 344	1634.71	1633.71	1633.79	-0.09	0	86	1	GVLNCFPDIFSTEK + Carbamidomethyl (C)
<input checked="" type="checkbox"/> 353	1717.86	1716.85	1716.92	-0.07	0	47	1	IGGIPIGIGQAFGGQPK
<input checked="" type="checkbox"/> 437	2132.01	2131.00	2131.09	-0.09	1	93	1	GVLNCFPDIFSTEKGAIQK + Carbamidomethyl (C)
<input checked="" type="checkbox"/> 439	2180.03	2179.03	2179.10	-0.07	0	129	1	GQIGLPPAAYVEVMSAAEAQK

<input checked="" type="checkbox"/>	443	2215.03	2214.03	2214.11	-0.08	0	26	1	VDAQVELGTFIHSMDVAVR
<input checked="" type="checkbox"/>	455	2291.11	2290.11	2290.20	-0.09	0	106	1	LSASGATSVQVQVDFPFASPLPR
<input checked="" type="checkbox"/>	498	2703.17	2702.16	2702.27	-0.12	0	154	1	AMYDPTGEPGSSSELSIATGDVLSVTR
<input checked="" type="checkbox"/>	499	2703.19	2702.18	2702.27	-0.09	0	(33)	1	AMYDPTGEPGSSSELSIATGDVLSVTR

5. [gi117137386](#) Mass: 51893 Total score: 521 Peptides matched: 8
Abelson Interacting Protein CG9749-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide	
<input checked="" type="checkbox"/>	297	1481.68	1480.67	1480.73	-0.06	1	36	1	QSLRDSYTNLER
<input checked="" type="checkbox"/>	360	1736.96	1735.95	1736.03	-0.09	1	93	1	IVAPINPEKPIKYVR
<input checked="" type="checkbox"/>	475	2444.18	2443.17	2443.27	-0.11	0	67	1	TSTGSQLAPIVPEQNLPGWVVK
<input checked="" type="checkbox"/>	496	2695.21	2694.21	2694.32	-0.11	0	(124)	1	KPIDYSMLDEIGHGINSAQHSQVR
<input checked="" type="checkbox"/>	497	2695.25	2694.25	2694.32	-0.07	0	170	1	KPIDYSMLDEIGHGINSAQHSQVR
<input checked="" type="checkbox"/>	505	2842.28	2841.27	2841.41	-0.14	1	77	1	TPFVVNPFQVPSHYAPNYPIGHKPR + Phospho (Y)
<input checked="" type="checkbox"/>	519	3075.54	3074.54	3074.61	-0.07	1	52	1	TSTGSQLAPIVPEQNLPGWVVKPIEK
<input checked="" type="checkbox"/>	521	3121.46	3120.46	3120.60	-0.15	0	40	1	GSSBGSVQSLLPSPVGGPPPTTKPTPPQMSK

6. [gi18084](#) Mass: 39533 Total score: 439 Peptides matched: 8
Hrb87F [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide	
<input type="checkbox"/>	12	631.27	630.26	630.31	-0.05	0	27	1	ARFEK
<input type="checkbox"/>	171	930.40	929.39	929.44	-0.05	1	16	3	AGGRGGQDR
<input checked="" type="checkbox"/>	180	955.43	954.42	954.47	-0.05	1	74	1	CGGGGGGGGR
<input checked="" type="checkbox"/>	199	1053.53	1052.52	1052.57	-0.05	0	36	1	LFIGGLDYR
<input checked="" type="checkbox"/>	221	1130.44	1129.44	1129.50	-0.07	0	70	1	GQGGGGGGQNR
<input checked="" type="checkbox"/>	238	1181.62	1180.61	1180.66	-0.05	1	75	1	RLFIGGLDYR
<input checked="" type="checkbox"/>	353	1700.66	1699.65	1699.75	-0.10	1	62	1	GGQDRGGGGGGGGQNR
<input checked="" type="checkbox"/>	465	2373.14	2372.13	2372.19	-0.06	1	25	1	GFATIEFDYDPVKIILQK

Proteins matching the same set of peptides:

- [gi18318](#) Mass: 39547 Total score: 439 Peptides matched: 8
P11 (hnRNP protein) [Drosophila melanogaster]
- [gi11036](#) Mass: 33707 Total score: 439 Peptides matched: 8
hrp36.1 [Drosophila melanogaster]
- [gi1829184](#) Mass: 39632 Total score: 439 Peptides matched: 8
heterogeneous nuclear ribonucleoprotein [Drosophila melanogaster]
- [gi117136622](#) Mass: 33650 Total score: 439 Peptides matched: 8
Heterogeneous nuclear ribonucleoprotein at 87F CG12745-PB, isoform B [Drosophila melanogaster]
- [gi148828065](#) Mass: 39476 Total score: 439 Peptides matched: 8
LD32727p [Drosophila melanogaster]

7. [gi11461109](#) Mass: 16965 Total score: 415 Peptides matched: 8
PDGF/VEGF receptor [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide	
<input type="checkbox"/>	221	1130.44	1129.44	1129.49	-0.06	0	11	9	GSNSFGVVQR + Phospho (ST)
<input checked="" type="checkbox"/>	266	1282.61	1281.61	1281.68	-0.08	0	67	1	YVNGHIVDIR
<input checked="" type="checkbox"/>	269	1308.60	1307.59	1307.65	-0.06	1	(25)	1	EFEPFRENLK
<input checked="" type="checkbox"/>	270	1308.61	1307.60	1307.65	-0.05	1	41	1	EFEPFRENLK
<input checked="" type="checkbox"/>	407	1974.95	1973.94	1974.03	-0.09	1	50	1	FNQALKQYVTPFSR
<input checked="" type="checkbox"/>	442	2202.98	2201.97	2202.06	-0.09	0	103	1	QSTDYLAALMGSPDELAPAAPR
<input checked="" type="checkbox"/>	459	2320.00	2318.99	2319.10	-0.10	0	78	1	NGMPTVDVADQAPEEIPMLR
<input checked="" type="checkbox"/>	483	2536.99	2535.98	2536.10	-0.12	0	66	1	LEGSSDFDFSSSETTFNFPGAR

Proteins matching the same set of peptides:

- [gi128874578](#) Mass: 165004 Total score: 415 Peptides matched: 8
PDGF- and VEGF-receptor related CG8222-PC, isoform C [Drosophila melanogaster]

8. [gi113430414](#) Mass: 41653 Total score: 402 Peptides matched: 7
actin E2 [Drosophila virilis]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>182</u>	976.38	975.37	975.44	-0.07	0	38	1	AGFAGDDAPR
<input checked="" type="checkbox"/> <u>215</u>	1198.65	1197.64	1197.70	-0.05	0	58	1	AVFPSIVGRPR
<input checked="" type="checkbox"/> <u>308</u>	1515.68	1514.67	1514.74	-0.07	0	(16)	1	IWHHTFYNELR
<input checked="" type="checkbox"/> <u>309</u>	1515.70	1514.69	1514.74	-0.05	0	66	1	IWHHTFYNELR
<input checked="" type="checkbox"/> <u>369</u>	1790.81	1789.80	1789.88	-0.08	0	108	1	SYELPDGQVITIGNR
<input checked="" type="checkbox"/> <u>404</u>	1959.98	1952.97	1953.06	-0.09	0	126	1	VAPSEHPVLLTEAPLNPK
<u>504</u>	2798.22	2797.21	2797.17	0.04	1	9	9	CDDDAGALVIDNGSGMCKAGFAGDDAPR + Carbamidometh

Proteins matching the same set of peptides:
gi117975545 Mass: 41673 Total score: 402 Peptides matched: 7
 Actin 88F CGS178-PA [Drosophila melanogaster]
gi1156773 Mass: 41748 Total score: 401 Peptides matched: 7
 actin

9. gi121587461 Mass: 149165 Total score: 385 Peptides matched: 10
 specifically Racl-associated protein 1 CG4931-PA [Drosophila melanogaster]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>148</u>	822.45	821.45	821.49	-0.04	0	58	1	HVQLLGR
<u>240</u>	1187.66	1186.65	1186.45	0.21	0	11	2	TELFQSFPR + 2 Phospho (ST)
<input checked="" type="checkbox"/> <u>268</u>	1300.68	1299.67	1299.75	-0.08	1	71	1	SIDLNKLITQR
<u>269</u>	1308.60	1307.59	1307.55	0.04	0	(14)	2	EFYLEMTMGR + 2 Oxidation (M)
<u>270</u>	1308.61	1307.60	1307.55	0.05	0	17	2	EFYLEMTMGR + 2 Oxidation (M)
<input checked="" type="checkbox"/> <u>322</u>	1564.74	1563.74	1563.81	-0.07	0	130	1	LNVGSPSTQLYMR
<u>345</u>	1650.85	1649.84	1649.91	-0.08	0	8	3	FQVLNSQIFSIILNK
<input checked="" type="checkbox"/> <u>433</u>	2115.91	2114.90	2115.00	-0.10	1	26	1	ANFDTNFEDRNGFVTIGIAK
<input checked="" type="checkbox"/> <u>477</u>	2464.21	2463.20	2463.30	-0.09	1	18	1	NLEVVPFLFGDMQIAPFNYYIKR
<input checked="" type="checkbox"/> <u>500</u>	2731.10	2730.09	2730.25	-0.15	1	48	1	LLEPTDHHQNKCECPVEAEBYER + Carbamidomethyl (C)

10. gi1158739 Mass: 50119 Total score: 288 Peptides matched: 8
 beta-1 tubulin
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>114</u>	738.31	737.31	737.35	-0.04	0	11	1	GSQQYR
<input checked="" type="checkbox"/> <u>334</u>	1601.73	1600.72	1600.81	-0.09	0	44	1	AVLVLEPGTMDSVR
<input checked="" type="checkbox"/> <u>339</u>	1620.76	1619.75	1619.83	-0.08	0	60	1	LEFFMPGFAPLTSR
<input checked="" type="checkbox"/> <u>352</u>	1691.79	1690.78	1690.86	-0.08	0	84	1	ALTVPELTQQMFDK
<input checked="" type="checkbox"/> <u>378</u>	1829.82	1828.81	1828.91	-0.10	1	41	1	INVYYNEASGCKYVPR
<input checked="" type="checkbox"/> <u>412</u>	2013.96	2012.95	2013.04	-0.09	1	26	1	MSATFIGNSTAIQELPKR
<input checked="" type="checkbox"/> <u>428</u>	2087.01	2086.00	2086.07	-0.07	1	10	1	GRYTEGAEVLVDSVLDVVRK
<input checked="" type="checkbox"/> <u>504</u>	2758.22	2757.21	2757.34	-0.12	0	13	1	SGPFGQIFRDPNFVFGSGGNNMRR

Proteins matching the same set of peptides:
gi124688737 Mass: 50115 Total score: 288 Peptides matched: 8
 -Tubulin at 56D CG9277-PB, isoform B [Drosophila melanogaster]
gi124688741 Mass: 51264 Total score: 288 Peptides matched: 8
 -Tubulin at 56D CG9277-PA, isoform A [Drosophila melanogaster]

11. gi117426899 Mass: 134357 Total score: 273 Peptides matched: 5
 Eps-15 protein [Drosophila melanogaster]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<u>152</u>	840.41	839.41	839.42	-0.01	0	31	4	FEEIFR
<input checked="" type="checkbox"/> <u>311</u>	1519.76	1518.75	1518.81	-0.06	0	43	1	SLOSELDTLTATLK
<input checked="" type="checkbox"/> <u>324</u>	1567.75	1566.74	1566.80	-0.06	0	102	1	AVVSVFDAFGEIGTR
<input checked="" type="checkbox"/> <u>416</u>	2023.07	2022.07	2022.13	-0.07	0	98	1	DIAPPAIIPPLVAVPEMTR
<input checked="" type="checkbox"/> <u>417</u>	2023.09	2022.09	2022.13	-0.05	0	(46)	1	DIAPPAIIPPLVAVPEMTR

12. gi1498894 Mass: 69658 Total score: 256 Peptides matched: 3
 poly(A)-binding protein [Drosophila melanogaster]
 Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>264</u>	1271.55	1270.54	1270.61	-0.07	0	63	1	FFGSQVATQMR
<input checked="" type="checkbox"/> <u>488</u>	2578.17	2577.16	2577.31	-0.14	0	181	1	AITGQQTAAPNMQIPIGAGIAGGAQQR
<input checked="" type="checkbox"/> <u>508</u>	2892.35	2891.34	2891.44	-0.10	0	13	1	MQQLGQIYQPNMAASGFPVETLPSNQR

Proteins matching the same set of peptides:
gi|21428474 Mass: 69832 Total score: 256 Peptides matched: 3
 LD24412p [Drosophila melanogaster]

13. gi|3869204 Mass: 61038 Total score: 226 Peptides matched: 5
 Cortactin [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>95</u>	705.31	704.30	704.35	-0.05	0	35	1	TIDGSGR
<input checked="" type="checkbox"/> <u>177</u>	947.38	946.37	946.43	-0.05	0	54	1	IDDGWWR
<input checked="" type="checkbox"/> <u>255</u>	1227.55	1226.54	1226.63	-0.08	1	45	1	TSTBAPPPKGSR
<input checked="" type="checkbox"/> <u>307</u>	1507.81	1506.81	1506.89	-0.08	0	77	1	VKFEVIEGAKPSNLR
<input checked="" type="checkbox"/> <u>418</u>	2025.95	2024.94	2025.01	-0.06	1	15	1	NRVGLFPANVYQVVQGNLS

Proteins matching the same set of peptides:
gi|24643611 Mass: 61046 Total score: 226 Peptides matched: 5
 Cortactin CG3637-PA [Drosophila melanogaster]

14. gi|24646107 Mass: 61359 Total score: 195 Peptides matched: 5
 CG6946-PB, isoform B [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>152</u>	840.41	839.41	839.45	-0.05	0	45	1	YIEIFR
<input checked="" type="checkbox"/> <u>234</u>	1168.55	1167.54	1167.62	-0.08	0	60	1	YIEVFTATEPK
<input checked="" type="checkbox"/> <u>278</u>	1345.69	1344.68	1344.73	-0.05	0	11	2	FFPEIRPANVR
<input checked="" type="checkbox"/> <u>347</u>	1657.74	1656.73	1656.83	-0.10	0	16	6	ATGAGGGVGGRRPGPYDIR
<input checked="" type="checkbox"/> <u>453</u>	2287.84	2286.83	2286.95	-0.12	0	61	1	GLNGEHEGGGNGGGMCNNGGQNSR

15. gi|24689266 Mass: 147385 Total score: 183 Peptides matched: 4
 Fak-like tyrosine kinase CG3969-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>235</u>	1482.80	1481.79	1481.87	-0.08	1	23	1	KYPLIIPANGLQR
<input checked="" type="checkbox"/> <u>306</u>	1507.91	1506.80	1506.88	-0.07	0	48	1	ALLDIGESPAATLLR
<input checked="" type="checkbox"/> <u>421</u>	2046.99	2045.98	2046.05	-0.07	1	35	1	CGDIISVLDNRNTGTPPWK
<input checked="" type="checkbox"/> <u>445</u>	2258.02	2257.01	2257.10	-0.09	0	79	1	TGYFNPSNTVAFLEGLPSSTR

16. gi|4377461 Mass: 95450 Total score: 166 Peptides matched: 4
 Dos protein [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>361</u>	1748.74	1747.73	1747.82	-0.09	0	90	1	NAHSAFIEESYDIPR
<input checked="" type="checkbox"/> <u>387</u>	1874.95	1873.94	1874.04	-0.09	1	31	1	LGAQAQLQQPIGPPSVDRK
<input checked="" type="checkbox"/> <u>397</u>	1926.85	1925.84	1925.97	-0.13	0	20	1	DLPQLSDTENTSPAIVAR
<input checked="" type="checkbox"/> <u>510</u>	2895.47	2894.46	2894.55	-0.10	0	28	1	FLPGVPLPGADLAIPNNPTPLNLDPK

Proteins matching the same set of peptides:
gi|24686127 Mass: 95476 Total score: 166 Peptides matched: 4
 daughter of sevenless CG1044-PA, isoform A [Drosophila melanogaster]

17. gi|38752635 Mass: 50275 Total score: 138 Peptides matched: 4
 LP10071p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>272</u>	1315.70	1314.69	1314.77	-0.08	1	40	1	ALRLPLQDVYK
<input checked="" type="checkbox"/> <u>396</u>	1920.01	1919.00	1919.07	-0.07	0	22	1	TLIDALDAILPPARPTEK
<input checked="" type="checkbox"/> <u>489</u>	2591.32	2590.32	2590.36	-0.04	1	34	1	EGRADGKTLIDALDAILPPARPTEK

517 3011.45 3010.44 3010.56 -0.12 0 43 1 SGDAAIIVNLVPSKPLCVFAPOEFPFLGR + Carbamidometh;

18. gi121423576 Mass: 129297 Total score: 128 Peptides matched: 4
LD4349Sp [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>256</u>	1237.58	1236.57	1236.78	-0.21	0	11	LSLVGNFAILLK
<input checked="" type="checkbox"/> <u>258</u>	1239.57	1238.56	1238.64	-0.08	0	64	LAEEFIPHQK
<input checked="" type="checkbox"/> <u>328</u>	1579.79	1578.78	1578.83	-0.05	0	29	SLSVVNIFLEEMK
<input checked="" type="checkbox"/> <u>419</u>	2037.94	2036.94	2036.98	-0.05	0	26	YYVQYLSGPDATLNIK

19. gi117647889 Mass: 15275 Total score: 117 Peptides matched: 2
Ribosomal protein S17 CG3922-PB [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>194</u>	1027.50	1026.49	1026.56	-0.07	0	62	LLEPFNIK
<input checked="" type="checkbox"/> <u>321</u>	1559.71	1558.71	1558.79	-0.08	0	55	GLQLTQENTNFKR

20. gi11498137 Mass: 148257 Total score: 110 Peptides matched: 5
Tama [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>187</u>	796.38	795.37	795.41	-0.05	1	11	GNEFRR
<input checked="" type="checkbox"/> <u>176</u>	942.41	941.40	941.46	-0.06	0	42	GSAFELYR
<input checked="" type="checkbox"/> <u>355</u>	1709.78	1708.77	1708.87	-0.10	0	12	GMPPIGNLPEETEK
<input checked="" type="checkbox"/> <u>441</u>	2196.07	2195.06	2195.14	-0.08	0	13	GPIMDGVSLQQLDRPVETPR + Oxidation (M)
<input checked="" type="checkbox"/> <u>469</u>	2404.04	2403.03	2403.10	-0.07	1	35	ENWDDVVVFSISIKFPAYER

21. gi1157658 Mass: 72190 Total score: 107 Peptides matched: 2
heat shock protein cognate 72

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>301</u>	1490.71	1489.70	1489.78	-0.08	0	60	VFAPEEISAMVLGK
<input checked="" type="checkbox"/> <u>339</u>	1887.89	1886.88	1886.96	-0.08	0	47	VTHAVVTVPAYFNDQKR

Proteins matching the same set of peptides:

gi124644408 Mass: 72216 Total score: 107 Peptides matched: 2
Heat shock protein cognate 3 CG4147-PD, isoform D [Drosophila melanogaster]

22. gi18072217 Mass: 221987 Total score: 107 Peptides matched: 5
Dscam [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>141</u>	807.37	806.37	806.41	-0.05	1	13	ANRYQR
<u>176</u>	942.41	941.40	941.26	0.14	0	9	TTEEMR + Oxidation (M); 2 Phospho (ST)
<input checked="" type="checkbox"/> <u>254</u>	1226.65	1225.64	1225.69	-0.05	1	24	LGGRFDFPVIR
<input checked="" type="checkbox"/> <u>300</u>	1483.69	1482.68	1482.77	-0.08	0	54	ASGNMPPBIIWIK
<input checked="" type="checkbox"/> <u>435</u>	2125.05	2124.04	2124.07	-0.02	1	10	VGWVSPPLESANGVIRTYK + Phospho (Y)

23. gi111038 Mass: 34982 Total score: 95 Peptides matched: 2
hrp40.1 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>466</u>	2377.05	2376.05	2376.15	-0.10	1	78	TYFGQFGNIVEVEMFEBKQK
<input checked="" type="checkbox"/> <u>513</u>	2937.43	2936.43	2936.49	-0.07	1	17	GFAPIVFTNTEADIKVSADEHIINSK

Proteins matching the same set of peptides:

gi111040 Mass: 36219 Total score: 95 Peptides matched: 2
hrp40.2 [Drosophila melanogaster]

24. [gi13255957](#) Mass: 53362 Total score: 51 Peptides matched: 1
EG:132E8.1 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 320	1837.88	1836.87	1836.95	-0.08	0	91	1	ALNNTVPEGGSQPIWVR

25. [gi13934641](#) Mass: 57182 Total score: 76 Peptides matched: 2
Drongo [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 190	995.45	994.44	994.49	-0.05	0	54	1	AVECQEQR
<input checked="" type="checkbox"/> 326	1873.86	1872.85	1872.89	-0.04	0	22	1	SISMATFTQDEIDFLR

Proteins matching the same set of peptides:

[gi117137344](#) Mass: 65482 Total score: 76 Peptides matched: 2
drongo CG3365-PB, isoform B [Drosophila melanogaster]

26. [gi124663131](#) Mass: 38158 Total score: 74 Peptides matched: 5
epsilon schachtel CG5654-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 212	1106.53	1105.52	1105.59	-0.06	1	29	1	NFRPVMKK
213	1105.48	1108.47	1108.54	-0.07	0	15	6	NFNNGPPPPR
318	1554.62	1553.61	1553.74	-0.13	0	7	3	GNEARNVTGSPGEPVR
320	1554.65	1553.64	1553.74	-0.11	0	(4)	7	GNEARNVTGSPGEPVR
<input checked="" type="checkbox"/> 328	1857.82	1856.81	1856.91	-0.10	1	24	1	NDTREDVFEHQSAIAR

27. [gi117647885](#) Mass: 21016 Total score: 72 Peptides matched: 2
Ribosomal protein L18A CG6510-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
177	947.38	946.37	946.53	-0.16	0	19	8	ITGEIVSIK
<input checked="" type="checkbox"/> 191	1000.58	999.57	999.62	-0.05	1	53	1	IKFPLVQR

28. [gi118476](#) Mass: 11504 Total score: 71 Peptides matched: 1
unnamed protein product [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 370	1790.86	1789.85	1789.90	-0.05	0	71	1	AARVEVEPYWGLEAK

Proteins matching the same set of peptides:

[gi117136320](#) Mass: 11506 Total score: 71 Peptides matched: 1
Ribosomal protein LPI CG4087-PA [Drosophila melanogaster]

29. [gi118859095](#) Mass: 7472 Total score: 68 Peptides matched: 1
Ribosomal protein S28b CG2998-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 282	1374.65	1373.64	1373.70	-0.06	0	68	1	BGDILTLLSESR

30. [gi124645526](#) Mass: 91432 Total score: 67 Peptides matched: 4
CG8516-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
28	705.31	704.30	704.33	-0.03	0	11	5	QMVDGR
317	1544.71	1543.71	1543.68	0.03	1	7	10	SDCYAANMLQKER + Oxidation (M)
<input checked="" type="checkbox"/> 430	2095.95	2094.95	2094.99	-0.04	0	26	1	SGAAPTPEQQLLVDDYFHV
<input checked="" type="checkbox"/> 476	2452.20	2451.19	2451.30	-0.10	1	24	1	KQQDGLGVSPLESAALLVDAQK

31. [gi124648551](#) Mass: 14576 Total score: 66 Peptides matched: 1
Ribosomal protein S30 CG15697-PB, isoform B [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 236	1170.54	1169.54	1169.60	-0.06	0	66	1	FVNFVQGFGR

32. [gi124680098](#) Mass: 44117 Total score: 63 Peptides matched: 2
quaking related 58E-2 CG5821-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 426	2077.93	2076.92	2077.02	-0.11	0	33	1	SHNNAYHCQPKPYVPAQR
<input checked="" type="checkbox"/> 531	3470.36	3469.35	3469.50	-0.15	0	31	1	QLQQQSNAAASGAGGGGGGGNGNGGAAGSGSNNGNGNR

33. [gi16690788](#) Mass: 594227 Total score: 60 Peptides matched: 5
Short stop/Kakapo long isoform [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 317	1544.71	1543.71	1543.81	-0.11	1	20	1	RMKPVGSELDQIR + Oxidation (M)
343	1631.77	1630.76	1630.87	-0.10	0	6	3	LPHVLEPLSSAESPIR
360	1736.96	1735.95	1735.75	0.20	0	19	2	QLLAADYGSDDLPSVK + 2 Phospho (ST)
387	1874.95	1873.94	1874.11	-0.17	1	15	3	QDVILIKNLLVSVQHR
408	1985.06	1984.05	1983.81	0.24	1	5	8	SLSDLTRLPSQADSVR + 3 Phospho (ST)

Proteins matching the same set of peptides:

[gi124653487](#) Mass: 614859 Total score: 60 Peptides matched: 5
short stop CG18076-PG, isoform G [Drosophila melanogaster]

34. [gi138047863](#) Mass: 17662 Total score: 59 Peptides matched: 3
similar to Drosophila melanogaster CG3195 [Drosophila yakuba]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
236	1419.68	1418.67	1418.70	-0.03	1	15	9	IGDIAKATSDWK
328	1579.75	1578.78	1578.64	0.14	1	(7)	5	IGDIAKATSDWK + 2 Phospho (ST)
<input checked="" type="checkbox"/> 337	1612.80	1611.79	1611.84	-0.05	0	45	1	BSGNIGFEDILAIAR

35. [gi13337433](#) Mass: 88776 Total score: 57 Peptides matched: 2
transitional endoplasmic reticulum ATPase TER94 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
194	1027.50	1026.49	1026.62	-0.13	1	20	4	EAILKANLR
<input checked="" type="checkbox"/> 338	1882.84	1881.84	1881.92	-0.08	0	37	1	SAAPCVLFFPELELSIAK + Carbamidomethyl (C)

Proteins matching the same set of peptides:

[gi117137560](#) Mass: 88803 Total score: 57 Peptides matched: 2
TER94 CG2331-PA, isoform A [Drosophila melanogaster]

[gi124682279](#) Mass: 32476 Total score: 57 Peptides matched: 2
TER94 CG2331-PB, isoform B [Drosophila melanogaster]

36. [gi124646932](#) Mass: 79184 Total score: 53 Peptides matched: 1
CG6904-PC, isoform C [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 501	2763.21	2762.20	2762.33	-0.12	1	54	1	MFDTCIQGNIFNADDLLQKDDLVK + Carbamidomethyl (C)

37. [gi120129705](#) Mass: 18464 Total score: 53 Peptides matched: 1
Ribosomal protein L21 CG12775-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 394	1913.99	1912.99	1913.05	-0.07	1	53	1	KLESPIALAPIPYEFIA

38. [gi112958781](#) Mass: 175184 Total score: 51 Peptides matched: 3
no-mechanoreceptor potential A long isoform [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
25	645.23	644.23	644.27	-0.05	0	25	1	YFCR + Carbamidomethyl (C)
<input checked="" type="checkbox"/> 290	1452.72	1451.71	1451.67	0.04	0	17	1	YPPAPIAFTGSSK + Phospho (ST)
463	2353.18	2352.17	2352.04	0.13	1	10	8	TNMSMFGSLHKTFTATGNLAR + 2 Oxidation (M); Pho

Proteins matching the same set of peptides:

[gi112958783](#) Mass: 174165 Total score: 51 Peptides matched: 3
no-mechanoreceptor potential A short isoform [Drosophila melanogaster]

39. [gi124640206](#) Mass: 21579 Total score: 51 Peptides matched: 1
Ribosomal protein L17 CG3203-PC, isoform C [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 273	1317.72	1316.71	1316.78	-0.07	1	51	1	KSABFLQLLR

40. [gi1755821](#) Mass: 72004 Total score: 51 Peptides matched: 1
ena polypeptide

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 446	2234.01	2233.01	2233.10	-0.10	1	51	1	KBDPQADLMGSLASQLQOFK

41. [gi121356037](#) Mass: 67309 Total score: 51 Peptides matched: 2
CG9796-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 207	1089.53	1088.52	1088.63	-0.12	0	34	1	IVFCNASLIR
322	1564.74	1563.74	1563.79	-0.06	1	17	3	LDRAYHSILLGRR + Phospho (ST)

42. [gi115010476](#) Mass: 155945 Total score: 50 Peptides matched: 1
GH06265p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 236	1419.68	1418.67	1418.74	-0.07	0	50	1	ITAAELEEELFQR

43. [gi124653903](#) Mass: 59130 Total score: 49 Peptides matched: 1
fusilli CG8205-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
152	840.41	839.41	839.46	-0.05	0	49	1	YIELFR

Proteins matching the same set of peptides:

[gi124653908](#) Mass: 102679 Total score: 49 Peptides matched: 1
fusilli CG8205-PF, isoform F [Drosophila melanogaster]

44. [gi117647697](#) Mass: 25418 Total score: 49 Peptides matched: 1
Ribosomal protein S5a CG8922-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 343	1631.77	1630.76	1630.82	-0.06	0	49	1	TIRECLADEELINARK + Carbamidomethyl (C)

45. [gi118860817](#) Mass: 39643 Total score: 47 Peptides matched: 1
Mec2 CG7635-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> 436	2125.99	2124.98	2125.07	-0.09	1	47	1	QSPKSNIVLDALDANFK

46. [gi116197837](#) Mass: 117317 Total score: 47 Peptides matched: 3
GH10594p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> 187	992.50	991.80	991.41	0.08	1	25	RLVCYCR + Phospho (Y)
<input checked="" type="checkbox"/> 188	992.55	991.54	991.41	0.13	1	(25)	RLVCYCR + Phospho (Y)
273	1317.72	1316.71	1316.78	-0.07	1	18	KPEKGITYLIR

47. [gi11228974](#) Mass: 106445 Total score: 47 Peptides matched: 3
vinculin [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
282	1374.65	1373.64	1373.68	-0.03	0	15	QVIDNATEISER
327	1612.80	1611.79	1611.80	-0.01	0	16	TMEHANQPILEARR + Phospho (ST)
<input checked="" type="checkbox"/> 356	1714.70	1713.69	1713.75	-0.06	1	17	GAEBARENRYLAAR + Phospho (Y)

Proteins matching the same set of peptides:

[gi12827490](#) Total score: 47 Peptides matched: 3

48. [gi18250181](#) Mass: 1840387 Total score: 45 Peptides matched: 4
D-Titin [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
145	822.45	821.45	821.50	-0.05	1	17	TYKIGIK
<input checked="" type="checkbox"/> 196	1052.49	1051.49	1051.48	0.00	0	19	VELEDYR
342	1627.80	1626.79	1626.63	0.16	0	8	SQEGTYEVIATNR + 2 Phospho (ST)
469	2404.04	2403.03	2403.07	-0.04	1	8	STLLIENFQSKFVGPFTCR + Carbamidomethyl (C); 2

49. [gi17576913](#) Mass: 77239 Total score: 45 Peptides matched: 2
E protein transcript B [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> 411	2002.08	2001.07	2001.14	-0.07	0	16	QILQILPPLQPPPYTOR
<input checked="" type="checkbox"/> 424	2066.96	2065.95	2066.03	-0.08	0	30	QALNSPDPFHPNLTSTR

50. [gi13056723](#) Mass: 183329 Total score: 44 Peptides matched: 4
translation initiation factor eIF4G [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> 252	1219.60	1218.59	1218.49	0.10	0	10	SNESAGNYIGK + Phospho (Y)
<input checked="" type="checkbox"/> 377	1825.92	1824.91	1824.99	-0.08	0	23	NVSILPQPNLMPSPFIR
459	2320.00	2318.95	2319.17	-0.13	1	10	FLKTYLTYCTQEVGPNFAR + Carbamidomethyl (C)
512	2910.36	2909.35	2909.15	0.20	1	5	VSAKISSIINYNEGQWSPNNESGK + 4 Phospho (ST)

51. [gi13444](#) Mass: 62291 Total score: 43 Peptides matched: 2
unnamed protein product [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
135	787.35	786.35	786.38	-0.04	1	20	RGGGGGNR
<input checked="" type="checkbox"/> 471	2413.09	2412.08	2412.18	-0.09	0	26	GQVNPICQDFSEVHLDPYVMK

Proteins matching the same set of peptides:

[gi12464463](#) Mass: 62791 Total score: 43 Peptides matched: 2

Rm62 CG10279-PE, isoform E [Drosophila melanogaster]

[gi145551833](#) Mass: 62929 Total score: 43 Peptides matched: 2

Rm62 CG10279-PB, isoform B [Drosophila melanogaster]

52. [gi13337431](#) Mass: 743005 Total score: 43 Peptides matched: 7
projectin [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>28</u>	648.23	644.23	644.27	-0.05	0	12	3	FYCR + Carbamidomethyl (C)
<u>90</u>	698.32	697.31	697.40	-0.09	1	11	2	RQAAPR
<input checked="" type="checkbox"/> <u>128</u>	757.45	756.45	756.31	0.14	1	2	1	EKGTDK + Phospho (ST)
<u>247</u>	1203.56	1202.56	1202.63	-0.08	1	3	2	RPKFGQWER
<u>418</u>	2022.99	2021.98	2021.84	0.14	0	2	7	STAPQVDVTGLSPGNEYK + 2 Phospho (ST)
<input checked="" type="checkbox"/> <u>423</u>	2057.97	2056.96	2056.77	0.19	0	18	1	LENDSDSNYNIDMESYR + Phospho (ST)
<u>501</u>	2763.21	2762.20	2762.19	0.01	1	6	3	VIAVNRGGPSDPSDPSSTIIICKPR + Carbamidomethyl (C)

53. gi117885887 Mass: 85029 Total score: 43 Peptides matched: 2
belle CG9748-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>211</u>	1103.62	1102.61	1102.66	-0.05	0	41	1	HAIPPIINGR
<u>402</u>	1944.78	1943.77	1943.85	-0.08	0	4	4	TSTNSVTGGVYVPPHLR + Phospho (ST); Phospho (Y)

54. gi119920688 Mass: 40476 Total score: 42 Peptides matched: 2
CG15436-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>177</u>	947.38	946.37	946.52	-0.15	0	24	5	LVNIFDAR
<u>337</u>	1612.80	1611.79	1611.70	0.08	0	19	2	TFAQQSTLQSHR + Phospho (ST)

55. gi1456984 Mass: 580269 Total score: 42 Peptides matched: 5
ryanodine receptor, calcium release channel [Drosophila melanogaster, Peptide, 5126 aa]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>188</u>	992.55	991.54	991.54	0.01	1	12	2	RAVIACFR + Carbamidomethyl (C)
<u>351</u>	1672.81	1671.80	1671.81	-0.00	0	3	9	GGNIDIQMGMLNHLK + 2 Oxidation (M)
<u>397</u>	1926.86	1925.84	1925.62	0.22	0	17	2	TEDMVTLSCTATGER + Carbamidomethyl (C); Oxidation (M)
<input checked="" type="checkbox"/> <u>420</u>	2041.83	2040.82	2040.85	-0.03	1	7	1	RGSVYIEDDYEMAEIR + Oxidation (M); Phospho (Y)
<u>458</u>	2308.15	2307.14	2307.19	-0.04	0	9	4	EAVSDFLVALTSQMPPAMLLK + Oxidation (M)

Proteins matching the same set of peptides:

gi1630870 Mass: 580203 Total score: 42 Peptides matched: 5
ryanodine receptor/calcium release channel - fruit fly [Drosophila melanogaster]
gi12160476 Mass: 580086 Total score: 42 Peptides matched: 5
ryanodine receptor homologue [Drosophila melanogaster]
gi12160477 Mass: 578745 Total score: 42 Peptides matched: 5
ryanodine receptor homologue [Drosophila melanogaster]
gi12160478 Mass: 578929 Total score: 42 Peptides matched: 5
ryanodine receptor homologue [Drosophila melanogaster]
gi117352465 Mass: 580642 Total score: 42 Peptides matched: 5
Ryanodine receptor 44F CG10844-PA, isoform A [Drosophila melanogaster]
gi117352467 Mass: 579301 Total score: 42 Peptides matched: 5
Ryanodine receptor 44F CG10844-PB, isoform B [Drosophila melanogaster]
gi117352468 Mass: 580458 Total score: 42 Peptides matched: 5
Ryanodine receptor 44F CG10844-PC, isoform C [Drosophila melanogaster]
gi117352471 Mass: 579118 Total score: 42 Peptides matched: 5
Ryanodine receptor 44F CG10844-PD, isoform D [Drosophila melanogaster]

56. gi124581733 Mass: 106066 Total score: 42 Peptides matched: 4
CG15627-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>153</u>	842.44	841.44	841.51	-0.08	1	(16)	5	DRLGCLR
<input checked="" type="checkbox"/> <u>154</u>	842.48	841.47	841.51	-0.04	1	23	1	DRLGCLR
<u>155</u>	842.50	841.49	841.51	-0.02	1	(14)	4	DRLGCLR
<u>198</u>	1052.49	1051.49	1051.53	-0.04	0	19	1	NLTADTVYR

57. gi14914494 Mass: 184034 Total score: 41 Peptides matched: 3
pacman protein [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
-------	----------	----------	----------	-------	------------	------	---------

<u>199</u>	1053.53	1052.52	1052.51	0.01	1	14	7	ISGSKTVCR
<u>206</u>	1073.42	1072.42	1072.50	-0.08	0	18	3	QNDCELIVR
<u>256</u>	1237.53	1236.57	1236.68	-0.11	0	10	6	GSDAQLLPALEK

58. [gi113857867](#) Mass: 100356 Total score: 40 Peptides matched: 2
CG10777-PB [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>463</u>	2353.18	2352.17	2352.25	-0.08	0	20	2	ELISVLEEAGQTFSSQALLDLAR
<input checked="" type="checkbox"/> <u>525</u>	3229.33	3228.32	3228.46	-0.14	0	20	1	SGPPTGSSYQQYAAGGNTYLQNGACCGVNWNR

59. [gi117593849](#) Mass: 176732 Total score: 39 Peptides matched: 2
no mechanoreceptor potential C CG11320-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>318</u>	1553.72	1552.71	1552.69	0.03	1	18	2	GVYLEMENKDGAR + Oxidation (M); Phospho (Y)
<input checked="" type="checkbox"/> <u>466</u>	2320.05	2319.04	2319.17	-0.13	1	25	1	SQTAVRLVSSRQTGTATNILR + Phospho (ST)

60. [gi18468](#) Mass: 48727 Total score: 39 Peptides matched: 2
unnamed protein product [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>246</u>	1200.67	1199.67	1199.48	0.19	0	13	1	SGQGAFGNMCR + Carbamidomethyl (C); Oxidation
<input checked="" type="checkbox"/> <u>274</u>	1326.62	1325.61	1325.68	-0.07	0	26	1	FVIWTESAFAR

Proteins matching the same set of peptides:

[gi124680747](#) Mass: 44998 Total score: 39 Peptides matched: 2
Ribosomal protein L4 CG5502-PA [Drosophila melanogaster]

61. [gi138048961](#) Mass: 6596 Total score: 39 Peptides matched: 1
similar to Drosophila melanogaster CG8495 [Drosophila yakuba]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>276</u>	1334.60	1333.59	1333.66	-0.06	0	39	1	GPATLWYSEPR

62. [gi117345862](#) Mass: 70248 Total score: 38 Peptides matched: 1
RE39037p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>286</u>	1479.72	1478.71	1478.79	-0.08	1	38	1	GIDALEKPFANFLR

63. [gi16225885](#) Mass: 180463 Total score: 38 Peptides matched: 3
defective chorion-1 fc177 protein precursor [Drosophila yakuba]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>207</u>	1089.53	1088.52	1088.43	0.09	1	16	10	SCRCSAGLR + Phospho (ST)
<u>318</u>	1553.72	1552.71	1552.64	0.07	0	16	8	QMTENPQMMQQR + 2 Oxidation (M)
<u>414</u>	2017.85	2016.85	2016.63	0.22	1	6	5	SGVSSGGSGTCQCKANR + 4 Phospho (ST)

64. [gi14056674](#) Mass: 213142 Total score: 39 Peptides matched: 2
plexin A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>273</u>	1317.72	1316.71	1316.64	0.08	1	31	2	YSLSEKILIR + Phospho (Y)
<u>386</u>	1873.86	1872.85	1872.93	-0.07	1	3	7	FLKECAGEPLMYLFR + Carbamidomethyl (C)

65. [gi124645017](#) Mass: 184141 Total score: 37 Peptides matched: 3

CG9626-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>12</u>	631.27	630.26	630.35	-0.09	0	13	8 AHLYK
<u>207</u>	1089.53	1088.52	1088.56	-0.05	0	13	6 DCQRALLSIR
<u>259</u>	1240.61	1239.60	1239.67	-0.07	1	9	2 QDNTLPRIQR

66. gi133049391 Mass: 17601 Total score: 37 Peptides matched: 1
similar to Drosophila melanogaster RpS18 [Drosophila yakuba]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>147</u>	831.38	830.37	830.42	-0.04	0	37	1 HYWGLR

67. gi124655594 Mass: 444454 Total score: 37 Peptides matched: 3
Dynein heavy chain at 62B CG15604-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>240</u>	1187.66	1186.65	1186.54	0.11	0	11	2 LVITPLTDR + 2 Phospho (ST)
<u>279</u>	1352.77	1351.76	1351.69	0.07	1	23	2 FASDDRYILPR
<input checked="" type="checkbox"/> <u>520</u>	3110.41	3109.40	3109.46	-0.07	1	3	1 RYEEVFSVEVFLNLALISLDDYNSTR + Phospho (Y)

68. gi162464380 Mass: 134481 Total score: 37 Peptides matched: 2
lingerer CG8715-PE, isoform B [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>163</u>	869.33	868.33	868.39	-0.06	0	14	6 SGGGGGGEAR
<input checked="" type="checkbox"/> <u>240</u>	1187.66	1186.65	1186.71	-0.06	0	24	1 VLLLLMTQR

69. gi1349665 Mass: 529831 Total score: 36 Peptides matched: 4
cytoplasmic dynein heavy chain

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>89</u>	688.37	687.36	687.36	0.00	0	15	3 ISSEER
<u>267</u>	1295.70	1294.70	1294.56	0.14	1	9	2 TSPLDDSPRK + Phospho (ST)
<u>397</u>	1926.85	1925.84	1925.82	0.02	0	11	5 GISESNLLAEFNFLLR + Phospho (ST)
<u>477</u>	2464.21	2463.20	2463.10	0.11	0	6	5 ENFINSIVSNEGTEITDDVR + Phospho (ST)

70. gi1433182 Mass: 199997 Total score: 36 Peptides matched: 3
receptor protein tyrosine phosphatase

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>217</u>	1121.46	1120.45	1120.43	0.03	0	3	5 IMSADSDFR + Phospho (ST)
<input checked="" type="checkbox"/> <u>318</u>	1553.72	1552.71	1552.59	0.13	1	29	1 QDRYEVHYQR + 2 Phospho (Y)
<u>406</u>	1965.06	1964.05	1963.96	0.10	0	5	7 VSITPDDAIQSVLYVER + Phospho (Y)

Proteins matching the same set of peptides:

gi124639796 Mass: 199951 Total score: 36 Peptides matched: 3
Protein tyrosine phosphatase 4E CG6899-PA, isoform A [Drosophila melanogaster]

71. gi124640994 Mass: 452309 Total score: 36 Peptides matched: 4
CG2989-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>188</u>	992.55	991.54	991.52	0.02	1	12	2 RSYVQPQR
<u>247</u>	1203.56	1202.56	1202.66	-0.11	1	3	2 RPYALNRSR
<u>347</u>	1657.74	1656.73	1656.82	-0.09	0	17	2 RFLSSITPSYISLR + Phospho (Y)
<u>376</u>	1813.90	1812.89	1812.92	-0.03	1	6	4 RPLSSITPSYISLR + Phospho (Y)

72. gi124560801 Mass: 61197 Total score: 36 Peptides matched: 2

CG5080-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>145</u>	822.45	821.45	821.43	0.02	1	27	2	RHAEFGK
<u>476</u>	2452.20	2451.15	2451.14	0.06	1	9	6	SAATEATERLQQAEANIQADVR + Phospho (ST)

73. gi116644386 Mass: 525785 Total score: 35 Peptides matched: 3
dynein heavy chain [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>240</u>	1187.66	1186.65	1186.54	0.11	0	11	2	LVITPLTDR + 2 Phospho (ST)
<u>273</u>	1317.72	1316.71	1316.72	-0.00	0	15	3	WTLAGVCLLLQT
<input checked="" type="checkbox"/> <u>325</u>	1579.79	1578.78	1578.67	0.11	1	13	1	RSLTFYDKPR + Phospho (ST); Phospho (Y)

74. gi117864174 Mass: 81077 Total score: 35 Peptides matched: 2
maverick CG1901-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>233</u>	1376.68	1375.67	1375.77	-0.10	1	5	1	RLETNQHPPIR
<input checked="" type="checkbox"/> <u>396</u>	1931.98	1930.97	1930.88	0.10	1	31	1	IMLLYSSSLATNFR + Phospho (ST); Phospho (Y)

Proteins matching the same set of peptides:

gi117844333 Mass: 80541 Total score: 35 Peptides matched: 2
REF9013p [Drosophila melanogaster]

75. gi113492045 Mass: 193232 Total score: 34 Peptides matched: 1
SPRINT-a [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>422</u>	2054.93	2053.92	2054.03	-0.11	1	34	1	VSQQQDSQQQOQPTKR

Proteins matching the same set of peptides:

gi113492047 Mass: 190231 Total score: 34 Peptides matched: 1
SPRINT-b [Drosophila melanogaster]

76. gi17692380 Mass: 121519 Total score: 34 Peptides matched: 2
gpl50 [Drosophila virilis]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>242</u>	1196.59	1195.58	1195.65	-0.06	1	28	1	VERYAPINQV
<u>427</u>	2077.93	2076.92	2076.88	0.04	0	7	2	YCTCSHNVNSYLVATCSR + Carbamidomethyl (C)

77. gi124584386 Mass: 250151 Total score: 34 Peptides matched: 4
crinkled CG7595-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<input checked="" type="checkbox"/> <u>245</u>	1206.52	1205.51	1205.65	-0.14	1	13	1	RIFPLSTQFR
<u>250</u>	1206.53	1205.53	1205.65	-0.13	1	(12)	4	RIFPLSTQFR
<u>273</u>	1317.72	1316.71	1316.75	-0.03	1	15	9	EEMAKLAALVFR
<u>440</u>	2185.99	2184.99	2185.06	-0.08	1	7	5	NYHVFYCIAGLSADEKSR

78. gi120151595 Mass: 28981 Total score: 34 Peptides matched: 2
LD24077p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
<u>351</u>	1672.91	1671.80	1672.00	-0.20	1	5	5	LIPAPRGITGIVSAPVPK
<input checked="" type="checkbox"/> <u>478</u>	2485.08	2484.07	2484.20	-0.13	0	29	1	EMPLGSTPYQAYSDFLSKPTPR

79. [gi16822326](#) Mass: 91133 Total score: 34 Peptides matched: 3
VAV protein [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
270	1308.61	1307.60	1307.65	-0.05	1	3	SGKIAFLLYLK + Phospho (Y)
356	1714.70	1713.69	1713.76	-0.07	0	17	QLSEFNWFAGNMDR
<input checked="" type="checkbox"/> 425	2073.99	2072.98	2072.81	0.17	1	17	DSHNLADYRVEQSHSR + 2 Phospho (ST)

Proteins matching the same set of peptides:

[gi124643216](#) Total score: 34 Peptides matched: 3

80. [gi124642261](#) Mass: 9505 Total score: 33 Peptides matched: 2
CG12375-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
190	898.45	894.44	894.47	-0.03	1	14	VQQTIRMK + Oxidation (M)
258	1239.57	1238.56	1238.55	0.02	1	20	TISKLAYQR + 2 Phospho (ST)

81. [gi113892027](#) Mass: 118042 Total score: 33 Peptides matched: 2
leucine rich repeat protein GP150 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
242	1196.59	1195.58	1195.65	-0.06	1	28	VHRYPINQV
363	1784.90	1783.90	1783.85	0.04	1	6	SNPKLQPIVDATITK + 2 Phospho (ST)

Proteins matching the same set of peptides:

[gi139752607](#) Mass: 118079 Total score: 33 Peptides matched: 2
RE46351p [Drosophila melanogaster]

82. [gi124654746](#) Mass: 38707 Total score: 33 Peptides matched: 2
CG1231-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
221	1130.44	1129.44	1129.59	-0.15	0	20	ANGDLALFPGR
<input checked="" type="checkbox"/> 408	1989.06	1984.05	1983.81	0.25	1	14	RCDTISVGRYQIGSLFR + Carbamidomethyl (C); 2 Ph

83. [gi12645435](#) Mass: 171922 Total score: 32 Peptides matched: 4
CHD3 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
206	1073.42	1072.42	1072.59	-0.17	1	13	IDGSIRGDLR
258	1239.57	1238.56	1238.66	-0.10	0	8	DRPAPTIDLNK
278	1345.69	1344.68	1344.71	-0.03	0	3	VGGNIEVLGFNAR
341	1622.79	1621.79	1621.94	-0.16	1	8	QHVLRIGVMSLIR

84. [gi15679048](#) Mass: 61351 Total score: 32 Peptides matched: 1
GMI3640p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
152	840.41	839.41	839.38	0.02	0	32	3	EMLER + Oxidation (M)

Proteins matching the same set of peptides:

[gi15685824](#) Total score: 32 Peptides matched: 1

85. [gi124652047](#) Mass: 104114 Total score: 32 Peptides matched: 2
hikaru genki CG2040-EB, isoform B [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide	
172	930.40	929.40	929.53	-0.13	0	12	9	TTGALSRRPK
207	1089.53	1088.52	1088.56	-0.05	1	20	3	LNKCLAEQCK + Carbamidomethyl (C)

Proteins matching the same set of peptides:

[gi124652049](#) Total score: 32 Peptides matched: 2

gi124682051 Total score: 32 Peptides matched: 2

86. gi12778363 Mass: 188990 Total score: 32 Peptides matched: 3
microtubule binding protein D-CLIP-190 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>81</u>	681.30	680.30	680.43	-0.14	1	20	1	KLPAPR
<u>838</u>	1603.73	1602.73	1602.72	0.00	1	11	10	PSLSECGIENLRR + Phospho (ST)
<u>438</u>	2171.86	2170.85	2170.97	-0.12	1	4	10	SNIPTATSGTGIPQPSKMK + 2 Phospho (ST)

Proteins matching the same set of peptides:

gi124584306 Mass: 188950 Total score: 32 Peptides matched: 3
CLIP-190 CG5020-PA, isoform A [Drosophila melanogaster]

87. gi15901820 Mass: 88414 Total score: 32 Peptides matched: 2
BcDNA.GH05095 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<u>238</u>	1181.62	1180.61	1180.50	0.11	1	11	6	RTASYANYR + Phospho (ST)
<u>277</u>	1343.61	1342.60	1342.68	-0.08	0	23	2	VSGGGGAGIGGQEVAK

Proteins matching the same set of peptides:

gi124684536 Total score: 32 Peptides matched: 2

88. gi124644634 Mass: 45578 Total score: 32 Peptides matched: 1
CG1137-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<u>194</u>	1027.50	1026.49	1026.62	-0.13	1	33	2	ILDQKALIR

89. gi124667715 Mass: 222966 Total score: 32 Peptides matched: 3
CG11451-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<u>104</u>	715.36	714.35	714.44	-0.09	1	11	2	RIAEVK
<u>224</u>	1139.53	1138.53	1138.61	-0.08	1	10	3	VVAHQENYSK
<u>302</u>	1492.78	1491.77	1491.70	0.07	0	12	2	EHFLNFHLSLR + Phospho (ST)

90. gi116768868 Mass: 45780 Total score: 32 Peptides matched: 1
LD09351p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<u>177</u>	947.38	946.37	946.47	-0.10	0	32	2	NAADSSALAK

Proteins matching the same set of peptides:

gi178214278 Total score: 32 Peptides matched: 1

91. gi17919 Mass: 94412 Total score: 31 Peptides matched: 3
unnamed protein product [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<u>257</u>	1239.56	1238.55	1238.62	-0.07	1	9	8	KEEIGTLLLEK + Phospho (ST)
<u>276</u>	1334.60	1333.59	1333.72	-0.12	0	11	8	CLYAAAITAKER + Carbamidomethyl (C)
<u>478</u>	2485.08	2484.07	2483.91	0.16	1	13	2	CIITIKSTRIISMYPEVEEK + Carbamidomethyl (C); Oxid

Proteins matching the same set of peptides:

gi1245885709 Total score: 31 Peptides matched: 3

92. gi14688227 Mass: 103752 Total score: 31 Peptides matched: 1
gamma-tubulin ring protein Dgrip91 [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>177</u>	947.38	946.37	946.56	-0.19	1	31	3 IMKLSLR + Oxidation (M)

Proteins matching the same set of peptides:
 gi124641556 Total score: 31 Peptides matched: 1

93. gi121357595 Mass: 71813 Total score: 31 Peptides matched: 2
 CG6654-PA [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>190</u>	995.45	994.44	994.54	-0.10	1	22	2 QHRLGIQR
<u>412</u>	2013.96	2012.95	2012.84	0.11	1	11	3 KHTGERPYACDLCPMR + Carbamidomethyl (C); Phospho

94. gi117844681 Mass: 25030 Total score: 31 Peptides matched: 1
 AT14762p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>177</u>	947.38	946.37	946.52	-0.14	1	31	4 IVNRSWR

95. gi11362653 Mass: 50322 Total score: 31 Peptides matched: 2
 tailless (tll) protein - fruit fly (Drosophila virilis)

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>1</u>	618.38	617.37	617.34	0.02	1	12	7 ACRLR
<u>191</u>	1000.58	999.57	999.46	0.11	0	19	2 VAMKNDAR + Oxidation (M)

Proteins matching the same set of peptides:
 gi12440921 Total score: 31 Peptides matched: 2
 gi166571248 Total score: 31 Peptides matched: 2

96. gi124582716 Mass: 313022 Total score: 31 Peptides matched: 4
 CG8486-PA, isoform A [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>98</u>	713.37	712.36	712.35	0.01	1	7	2 KASLSK + Phospho (ST)
<u>116</u>	743.38	742.37	742.43	-0.07	0	5	4 LLNDLR
<u>190</u>	995.45	994.44	994.49	-0.05	0	15	4 SIFCIEQR
<u>420</u>	2041.83	2040.82	2040.90	-0.08	0	6	2 QALTFIAGYDTEVRAVR + Phospho (ST); Phospho (Y)

97. gi121063933 Mass: 51817 Total score: 31 Peptides matched: 1
 AT01380p [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>152</u>	840.41	839.41	839.35	0.06	0	31	4 MEELMR + 2 Oxidation (M)

98. gi124583241 Mass: 113192 Total score: 30 Peptides matched: 3
 CG4839-PB, isoform B [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>265</u>	1273.53	1272.52	1272.60	-0.09	1	13	3 KIATLCCGAFGR + Phospho (ST)
<u>325</u>	1570.79	1569.79	1569.77	0.01	0	9	4 LDEQCCEQIVANR
<u>337</u>	1612.80	1611.79	1611.62	0.17	0	12	4 EMVINAMRPASYSR + Phospho (ST); Phospho (Y)

99. gi121357161 Mass: 29718 Total score: 30 Peptides matched: 1
 Ribosomal protein L6 CG11522-PB, isoform B [Drosophila melanogaster]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>429</u>	2090.97	2089.96	2090.06	-0.10	1	30	1 VDLGAFKYPEHLNDAYFR

Proteins matching the same set of peptides:

gi124651668 Mass: 27681 Total score: 30 Peptides matched: 1
 Ribosomal protein L6 CG11522-PA, isoform A (Drosophila melanogaster)

100. gi13403156 Mass: 154480 Total score: 30 Peptides matched: 2
 DPP2C1 (Drosophila melanogaster)

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<u>135</u>	787.35	786.35	786.37	-0.03	0	19	5 VGGGGGGGGGR
<u>322</u>	1564.74	1563.74	1563.66	0.07	1	12	8 VVPSSSSMRTR + Oxidation (M); Phospho (ST)

Proteins matching the same set of peptides:
 gi124639731 Total score: 30 Peptides matched: 2

Unassigned queries: (no details means no match)

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss Score	Rank	Peptide
<input checked="" type="checkbox"/> <u>145</u>	832.43	831.42	831.48	-0.06	1	29	1 FVRWPK
<input checked="" type="checkbox"/> <u>235</u>	1405.67	1404.66	1404.73	-0.07	0	29	1 TQLSFVVPMMR
<input checked="" type="checkbox"/> <u>275</u>	1352.77	1351.76	1351.82	-0.06	0	29	1 IAEALDFPLRLR
<input checked="" type="checkbox"/> <u>347</u>	1687.74	1656.73	1656.91	-0.18	1	28	1 RASASMVISINPALR
<input checked="" type="checkbox"/> <u>119</u>	742.38	742.37	742.41	-0.04	0	28	1 ILGFFF
<input checked="" type="checkbox"/> <u>197</u>	1044.52	1043.51	1043.50	0.02	0	27	1 DVAAQEQQR
<input checked="" type="checkbox"/> <u>431</u>	2097.01	2096.00	2096.09	-0.09	0	26	1 EIPQIQIQMPFGMALOR
<input checked="" type="checkbox"/> <u>351</u>	1672.81	1671.80	1671.88	-0.08	0	26	1 RPAVDLAFHNLIYR
<input checked="" type="checkbox"/> <u>163</u>	869.33	868.33	868.56	-0.23	1	26	1 RQLLIAR
<input checked="" type="checkbox"/> <u>345</u>	1650.88	1649.84	1649.72	0.11	1	25	1 QFNCPANLASHRR + Carbamidomethyl (C); Phospho
<input checked="" type="checkbox"/> <u>336</u>	1603.76	1602.75	1602.74	0.01	0	24	1 FSLNMLIHISIR + 2 Phospho (ST)
<input checked="" type="checkbox"/> <u>277</u>	1343.61	1342.60	1342.86	0.04	1	24	1 QKDKQSEHDEL
<input checked="" type="checkbox"/> <u>155</u>	842.50	841.45	841.51	-0.02	1	24	1 RSLILPR
<input checked="" type="checkbox"/> <u>164</u>	869.36	868.35	868.37	-0.02	1	23	1 YRHSAR + Phospho (Y)
<input checked="" type="checkbox"/> <u>340</u>	1621.86	1620.85	1620.94	-0.09	1	22	1 VELLDSWILKLR
<input checked="" type="checkbox"/> <u>325</u>	1570.75	1569.75	1569.82	-0.08	1	22	1 LYEKIDLILNLR + Phospho (Y)
<input checked="" type="checkbox"/> <u>239</u>	1447.74	1446.73	1446.65	0.08	1	22	1 ISSSPAKTLVER + 2 Phospho (ST)
<input checked="" type="checkbox"/> <u>171</u>	930.40	929.39	929.50	-0.11	1	22	1 RQQVISR
<input checked="" type="checkbox"/> <u>1</u>	618.36	617.37	617.34	0.02	1	21	1 VGMRR
<input checked="" type="checkbox"/> <u>463</u>	2353.18	2352.17	2351.94	0.23	1	21	1 LEQMSKTSAVIARSTITSSDR + 3 Phospho (ST)
<input checked="" type="checkbox"/> <u>135</u>	787.35	786.35	786.35	-0.01	1	21	1 RHYVR + Phospho (Y)
<input checked="" type="checkbox"/> <u>302</u>	1492.78	1491.77	1491.85	-0.08	0	20	1 IQQVLEPLQQQVR
<input checked="" type="checkbox"/> <u>39</u>	687.37	687.36	687.37	-0.01	0	20	1 LINGQR
<input checked="" type="checkbox"/> <u>410</u>	1995.92	1994.91	1994.82	0.10	1	20	1 VLSLSSPPGDPGITPKSAR + 3 Phospho (ST)
<input checked="" type="checkbox"/> <u>172</u>	930.40	929.40	929.50	-0.11	1	20	1 RQQVISR
<input checked="" type="checkbox"/> <u>335</u>	1603.73	1602.73	1602.65	-0.12	1	20	1 IAAKTEVWLLSSR + Phospho (ST)
<input checked="" type="checkbox"/> <u>303</u>	1495.75	1494.74	1494.74	0.01	1	19	1 GAPEGVLERCTEAR
<input checked="" type="checkbox"/> <u>295</u>	1464.64	1463.63	1463.81	-0.18	1	18	1 GPSVIEAPSLPKNR
<input checked="" type="checkbox"/> <u>77</u>	674.33	673.32	673.39	-0.06	1	18	1 RTIER
<input checked="" type="checkbox"/> <u>244</u>	1197.57	1196.56	1196.51	0.05	0	18	1 GSNWRGLSR + Phospho (ST)
<input checked="" type="checkbox"/> <u>170</u>	927.38	926.37	926.54	-0.17	1	18	1 GLIGRAGQR
<input checked="" type="checkbox"/> <u>292</u>	1453.77	1452.76	1452.70	0.07	1	17	1 IINGSVAKADETR + Phospho (ST)
<input checked="" type="checkbox"/> <u>231</u>	1354.64	1353.64	1353.68	-0.04	1	17	1 TIYSVTAKVHR + Phospho (ST)
<input checked="" type="checkbox"/> <u>141</u>	807.37	806.37	806.41	-0.04	1	17	1 RTMIDR + Oxidation (M)
<input checked="" type="checkbox"/> <u>355</u>	1918.88	1917.88	1917.74	0.14	1	17	1 FQWSEBRTLOSSMK + Oxidation (M); 2 Phospho (S
<input checked="" type="checkbox"/> <u>153</u>	842.44	841.44	841.54	-0.10	0	17	1 ATVLLGIR
<input checked="" type="checkbox"/> <u>55</u>	667.26	666.25	666.37	-0.12	0	16	1 YPPIK
<input checked="" type="checkbox"/> <u>324</u>	1854.80	1853.79	1853.72	0.07	1	16	1 RASCSAMDSQAVTER + Oxidation (M); 2 Phospho
<input checked="" type="checkbox"/> <u>234</u>	1463.69	1462.68	1462.62	0.06	1	16	1 TDSKVLOQAGTR + 2 Phospho (ST)
<input checked="" type="checkbox"/> <u>405</u>	1993.84	1992.83	1992.80	0.03	1	16	1 KEYSTNSSLAGQSILR + 3 Phospho (ST)
<input checked="" type="checkbox"/> <u>327</u>	1578.65	1577.64	1577.62	0.02	0	15	1 EMSANGASLDANFR + Oxidation (M); Phospho (ST)
<input checked="" type="checkbox"/> <u>433</u>	2665.02	2664.01	2663.84	0.17	1	15	1 SCISTVSSVSDSGGRTTATSGR + 6 Phospho (ST)
<input checked="" type="checkbox"/> <u>257</u>	1239.56	1238.55	1238.59	-0.05	0	15	1 HSWYHGPIER
<input checked="" type="checkbox"/> <u>262</u>	1245.64	1244.64	1244.56	0.08	1	15	1 EPTATIQK + 2 Phospho (ST)
<input checked="" type="checkbox"/> <u>265</u>	1273.53	1272.52	1272.67	-0.15	0	15	1 VFKPHTNFQR

332	1595.85	1594.84	1594.83	0.01	1	7	1	IFPGLSKVMSALPR + Phospho (ST)
316	1542.71	1541.70	1541.80	-0.09	1	7	1	RIEMAEALSLTER + Oxidation (M)
400	1941.82	1940.81	1940.66	0.15	0	7	1	QTCGLPHTMYIDR + Oxidation (M); 3 Phospho (S
162	868.49	867.48	867.36	0.12	1	7	1	SSSHKSR + Phospho (ST)
314	1537.60	1536.59	1536.69	-0.10	1	7	1	QNOQRWQNSIR + Phospho (ST)
516	3001.26	3000.25	3000.44	-0.19	0	7	1	CLDAWCCVAVGGDGLFHEIVNGLLQR + 2 Carbamidomet
375	1809.77	1808.76	1808.79	-0.03	0	7	1	MNSSCADILLFPAYK + Carbamidomethyl (C); Phosph
488	2556.97	2555.96	2556.06	-0.10	0	6	1	CLKNTFPGAACNLCPGPGYDAIK + 2 Carbamidomethyl
415	2022.99	2021.98	2021.78	0.20	1	6	1	KIIFSESSGGGSENYCSK + Phospho (ST); Phospho (
346	1653.82	1652.82	1652.80	0.01	1	6	1	MCPDKVTVLLVGNK + Carbamidomethyl (C); Phosph
462	2352.13	2351.12	2350.98	0.14	1	6	1	TFMVASQDDKSNFVSSLAK + Oxidation (M); 2 Phos
107	722.28	721.27	721.34	-0.07	0	6	1	SLSCQK + Carbamidomethyl (C)
372	1806.76	1805.75	1805.98	-0.22	1	6	1	KALYSVQEBDLGLTIK
431	2530.39	2529.39	2529.28	0.11	1	6	1	RQALPILVFGNELVEQFYSR + Phospho (Y)
36	653.07	652.06	652.31	-0.24	0	6	1	YLDDK
100	713.43	712.42	712.35	0.07	1	6	1	KGTIVK + Phospho (ST)
315	1541.80	1540.79	1540.82	-0.03	1	6	1	RRPHAVSDEIHPK
492	2647.21	2646.20	2646.34	-0.14	1	6	1	LIAAGANTVITLGLKLGAVFGSADSK + 2 Phospho (ST)
161	868.47	867.46	867.36	0.10	1	6	1	SSSHKSR + Phospho (ST)
528	3337.62	3336.61	3336.52	0.09	1	6	1	GIKMKLLATTVAALFTICFLVSCPMPK + 2 Carbamidome
527	3323.59	3322.59	3322.51	0.08	1	5	1	SLVLEIMCSNVSTTFIECPSECKEQLGIK + 2 Phospho (S
71	669.40	668.39	668.34	0.06	0	5	1	NVGSER
10	629.17	628.16	628.29	-0.13	0	5	1	TSMPK + Oxidation (M)
402	1944.74	1943.73	1943.84	-0.09	1	5	1	VPIKSEARSGSAYEDPR + Oxidation (M); Phospho (
452	2283.09	2282.08	2282.07	0.01	0	5	1	VBPQLADLEPVVTASSCFVR + Carbamidomethyl (C);
26	643.22	642.21	642.27	-0.06	0	5	1	HNNCR
381	1838.84	1837.83	1837.76	0.07	0	5	1	NNMSQYMDIAFTSER + 2 Oxidation (M)
503	2793.11	2792.10	2792.23	-0.13	0	5	1	CALQELDITATDLSTECVLDMLSR + Carbamidomethyl (
313	1535.59	1534.58	1534.61	-0.02	1	4	1	GPTNVSCYKCNR + 2 Carbamidomethyl (C); Phosph
451	2272.01	2271.00	2270.90	0.10	0	4	1	NGSLSEAQLLDTLPSLASK + 4 Phospho (ST)
445	2225.02	2224.02	2223.91	0.11	1	3	1	GSERMEVNADDLLEMYEK + Oxidation (M); Phospho
38	656.05	655.04	655.26	-0.22	0	0	1	CFDSK + Carbamidomethyl (C)
22	642.18	641.17	641.30	-0.13	0	0	1	YGSTSK
2	622.08	621.07						
4	625.40	624.40						
5	625.88	624.87						
6	626.20	625.19						
7	628.16	627.15						
8	628.16	628.15						
9	628.16	628.15						
11	630.58	629.57						
13	632.58	631.57						
14	632.60	631.59						
15	634.60	633.59						
16	635.06	634.05						
17	640.97	639.96						
18	640.98	639.97						
20	642.13	641.13						
21	642.18	641.17						
23	642.61	641.60						
24	642.99	641.98						
25	643.18	642.17						
27	644.02	643.01						
30	646.16	645.16						
31	646.63	645.63						
32	650.04	649.03						
33	650.94	649.93						
34	650.96	649.95						
35	652.94	651.93						
37	656.01	655.00						

Search Parameters

Type of search : MS/MS Ion Search
Enzyme : Trypsin
Variable modifications : Carbamidomethyl (C), Oxidation (M), Phospho (ST), Phospho (Y)
Mass values : MONOISOTOPIC
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 0.25 Da
Fragment Mass Tolerance : ± 0.25 Da
Max Missed Cleavages : 1
Instrument type : MALDI-TOF-TOF
Number of queries : 536

Mascot: <http://www.matrixscience.com/>

8.0 Supplementary Information
General Protocols and Stock Solutions

Table 8.0: SDS-PAGE mini-gels (BioRad). Refer to Table 8.1 for appropriate percentage gels

1. Ensure gel apparatus is thoroughly rinsed. Organize caster, glass plates, spacers and combs. Insert comb and mark 1 cm with a felt marker to mark the top position of the separating gel. Ensure apparatus is level.
2. Remove comb and prepare gel solutions.
3. Make 10% (w/w) ammonium persulfate (BioRad Cat # 161-0700) in 1 mL centrifuge tube. (e.g. ~20 mg into 200 μ L).
4. Prepare separating and stacking gels. Add TEMED just before addition to casting apparatus. Prepare each solution in 10 mL tube.
5. Cast separating gel. Add TEMED. Mix gently. Using a disposable dropper add separating gel solutions up to the mark made in step 1. Do not introduce bubbles.
6. Level. To level and remove any bubbles, add water (above the separating gel) till slight overflow is observed. Allow to gel to cure for 40 minutes or use any left over gel within the tube to test polymerization.
7. Remove water. Rinse with ethanol and dry with #1 (Whatman) 6 inch filter paper (folded and cut).
8. Cast stacking gel. Add TEMED. Mix gently. Using a disposable dropper add stacking gel solution till slight overflow. Add comb until flush and centered. Avoid trapping air. Top off with excess stacking solution as needed. Wait 10 minutes till gel is set.

9. Prepare loading buffer, LB (2x) without β -mercaptoethanol. Add β -mercaptoethanol (10% v/v) just prior to use.
 10. Mix sample with LB to 1x concentration. Boil 1-2 minutes.
 11. Set up 3 rinse beakers, (1 x MeOH and 2 x ddH₂O) to wash syringe or gel loader tips between sample loads.
 12. Add 1x LB (+ β -mercaptoethanol) to empty lanes.
 13. Run gels at 200V for 45 minutes or until dye front runs off the gel.
 14. Stain gel for MALDI-MS or prepare for western-blot.
-

Table 8.1a: SDS-PAGE Separating Gels	6%	8%	10%	12%	12.5%	15%
ddH ₂ O	5.3 mL	4.6 mL	4.0 mL	3.3 mL	3.126 mL	2.3 mL
30% Acrylamide + Bisacrylamide	2.0 mL	2.7 mL	3.3 mL	4.0 mL	4.17 mL	5.0 mL
1.5 M Tris (pH 8.8)	2.5 mL					
10% SDS	100 µL	10 µL				
(Fresh) 10 % (w/w) Ammonium Persulfate	100 µL	10 µL				
TEMED	8 µL	6 µL	4 µL	4 µL	8µL	8µL
Total Volume	10 mL					

Table 8.1b: SDS-PAGE Stacking Gel (4%)

ddH ₂ O	6.1mL
30% Acrylamide + Bisacrylamide	1.3 mL
0.5 M Tris (pH 6.8)	2.5 mL
10% SDS	100 µL
(Fresh) 10 % (w/w) Ammonium Persulfate	50 µL
TEMED	10 µL
Total Volume	10 mL

Table 8.2: Western-blot

Day 1.

1. Obtain ice pack from -80°C. Replace promptly.
2. Make transfer buffer.
3. Prepare transfer membranes. PVDF (5 x 8.5 cm)
4. Get glass tray, transfer buffer, stir bar, magnetic stirrer, scrub pads, filter papers, and transfer membranes.
5. PVDF needs to be 'wetted' using MeOH before placing in transfer buffer.
6. Remove stacking gel for the gel. Remove gel from glass.
7. Assemble transfer sandwich according. Clear portion is positive, black is negative side. Proteins will transfer to the positive side and onto membrane. (-ve, bottom), scrub pad, 2x filter paper, gel, transfer membrane, 2x filter, scrub pad, (+ve, top). Assemble and saturate scrub pads with transfer buffer. Avoid introducing air.
8. Place ice pack in transfer chamber along with stir bar.
9. Place assembled apparatus, noticing electrode orientation. Place on stirrer.
10. Fill chamber with transfer buffer.
11. Run transfer for 100 V for 45 minutes. Turn on stirrer.
12. Once complete disassemble apparatus. Block membrane with TBST with 5% skim milk powder for 1 hour on shaker or with rotation. (blocking buffer: 5g skim milk powder with 95 mL TBST)
13. Take 10 mL of blocking buffer and dilute with 40mL of TBST to make 1% skim milk solution. Add primary antibody to appropriate concentration. E.g. 1:1000 = 5 μ L:5mL TBST 1% skim milk.
14. Shake overnight at 4°C.

Day 2

15. Wash 3 x 10 minutes with TBST.
 16. Prepare 2° antibodies in 1% skim milk in TBST
 - i) monoclonal Sigma A-8924 MS HRP 1:3000
 - ii) polyclonal Sigma A-6154 Rb HRP 1:5000
-

-
17. Wash 3 x 10 minutes with TBST.
 18. Prepare development cases, film and ECL I + II. Take to dark room
 19. Mix ECL I + II. Dip membrane and swirl for 1 minute. Remove membrane and dry.
 20. Turn off lights. Turn on red light. Expose film for regular intervals (e.g. 30 sec, 1 min, 2min, 4 min, 8 min. Store exposed film in case.
 21. Develop film using automated developer. Determine if longer or shorter exposures are necessary.

Table 8.3: *In-gel* digestion and peptide recovery for MALDI-MS

1. Excise the spot/band of interest, cut into smaller 1mm² pieces.
 2. Shrink the gel pieces in acetonitrile for 10 minutes, shake. Remove excess ACN, and dry under using SpeedVac.
 3. Rehydrate in 100 mM NH₄HCO₃ for 10 minute, shake. Repeat step 2.
 4. Cover all gel pieces with 10 mM DTT in 100 mM NH₄HCO₃. Incubate for 1 h at 56°C.
 5. Cool to room temperature, remove DTT solution and add equal volume of 55 mM iodoacetamide in 100 mM NH₄HCO₃. Incubate for 45 min in dark. Remove supernatant.
 6. Shrink the gel with ACN for 10 minutes (shake) and remove.
 7. Swell gel pieces with 100 mM NH₄HCO₃ and shrink again.
 8. Remove excess liquid, add ACN, to dehydrate. Remove liquid. Dry in SpeedVac. Gel pieces should appear white (if stained with Coomassie blue)
 9. Swell gel pieces at 4°C for 45 min in buffer containing trypsin and 50 mM NH₄HCO₃ (approx. 5 µL/mm² gel). The gel pieces should just be covered. (Trypsin should be prepared at 12.5 ng/µL).
 10. Digest overnight at 37°C overnight (8-16 hours).
 11. Centrifuge, collect supernatant and set aside.
 12. Add 20 mM NH₄HCO₃, shake (10 minutes), centrifuge and extract adding to supernatant tube (step 11).
 13. Extract with 50 % (v/v) ACN, 5% (v/v) TFA. Shake 10 minutes. Centrifuge, collect and repeat 2X.
 14. Dry sample down to in SpeedVac.
 15. Store in Freezer and reconstitute with appropriate buffer for ZipTip™ or LC-MS.
-

Table 8.4: ZipTip™ and application to MALDI plate

Desalting

1. Obtain 10 μ L pipette. Resuspend sample in a minimum volume of 5% (v/v) trifluoroacetic acid (TFA), \sim 5 μ L.
2. Wash ZipTip (Millipore, Bedford, MA) with three pipette volumes (\sim 10 μ L) using 60% (v/v) acetonitrile (ACN) containing 5% (v/v) TFA. Aspirate each volume by pipetting in and out 3X.
3. Equilibrate tip with 5% (v/v) ACN, aspirate 3X.
4. Extract the analytes by pipetting and aspirating 10-20 times.
5. Wash the tip with a solution of 5% (v/v) TFA using 3 pipette volumes and discard onto Kimwipe.
6. Elute the analytes using 3-5 μ L of a solution of 60% (v/v) ACN and 5% (v/v) TFA.

Application to MALDI plate

1. Prepare fresh solution of 10mg DHB/1mL 40% (v/v) ACN.
 2. Take 0.5 – 1.0 μ L of sample (eluent in Step 6 of the desalting protocol) and pipette onto MALDI plate.
 3. Immediately add matrix solution, 1:1 volume and allow to air dry for 5-10 minutes.
 4. Repeat the process for remaining samples.
 5. Insert target in the MALDI-MS instrument
-

Table 8.5: Cell passage (HeLa, PC12)

(protocol provided by X. Li)

1. From cell culture plates, remove growth media with sterile suction pipette and discard in appropriate biohazard waste. Perform all steps under sterile conditions. Wipe all surfaces with (50%) ethanol solution.
 2. Add 2 mL of Trypsin EDTA (37°C, Gibco cat #25300-54, 0.05% trypsin, 0.53mM EDTA+4Na). Incubate 37°C for 5 minutes. Swirl.
 3. Add 5 mL (37°C) of Dulbecco's Modified Eagle Medium (cat # 11995-065) which has been supplemented with 12.5% Horse serum and 2.5% FBS 1% penicillin-streptomycin.
 4. Aspirate to dislodge adherent cells.
 5. Transfer half of the material (2.5 mL) to new dish.
 6. Check transfer under microscope, 36°C (CO₂) incubator,
 7. Periodically check cell confluence and change media when changes from pink to orange.
-

Table 8.6: Immunohistochemistry – Cultured Cells

DAY1

1. From cell culture plates, remove growth media with sterile suction pipette and discard in appropriate biohazard waste. Perform all steps under sterile conditions. Wipe all surfaces with (50%) ethanol solution.
 2. Add 2 mL of Trypsin EDTA (37°C, Gibco cat #25300-54, 0.05% trypsin, 0.53mM EDTA+4Na). Incubate 37°C for 5 minutes. Swirl.
 3. Add 5 mL (37°C) of Dulbecco's Modified Eagle Medium (cat # 11995-065) which has been supplemented with 12.5% Horse serum and 2.5% FBS (1% penicillin-streptomycin). Aspirate to dislodge cells.
 4. Transfer cells to poly L-lysine treated glass slides (4 well slide, BD falcon cat. # 354114). Distribute cells evenly. Check transfer under microscope.
 5. Incubate @ 36°C, (CO₂) to confluence. Check periodically.
 6. Remove media. Rinse with 2 x 1mL PBS solution.
 7. Fix cells with 1% paraformaldehyde for 5 minutes.*
- * Alternatively, cells maybe stored @ 4°C in fix solution for up to 2 days. Label and date.
8. Calculate primary antibody dilution required. Usually 1:50 to 1:500, check data sheet. Ideally this should be done well in advance. (For 1:100 dilution, add 8.3 µL of antibody to 800 µL 1.5 TBST + 40µL serum). Plan controls appropriately (e.g. primary antibody omission).
 9. Prepare 1.5 TBST. Prepare sufficient amounts for 200 µL/ well or 800µL/slide. Add 40 µL of blocking normal donkey serum (5%, NDS) or normal goat serum (5%, NGS) as appropriate (not to interfere with 2° antibody recognition). Add primary antibody amounts as determined in step 8.
 10. Add 200 µL/well of primary anti-body (1.5 TBST, 5% NDS/NGS) solution prepared in step 8 & 9.
 11. Incubate overnight at 4°C. Carefully label and record. Continued...
-

Table 8.6: Immunohistochemistry – Cultured Cells continued.

Perform secondary antibody calculations, see step 12.

Day 2

12. Calculate 1.5 TBST volumes based on # slides (800 μ L/slide) or # wells (200 μ L/well). Add appropriate amount of blocking serum (NDS or NGS), e.g. 40 μ L/800 μ L 1.5 TBST or 10 μ L/200 μ L 1.5 TBST. Add appropriate amount of secondary antibody, see step 13.

13. Secondary antibodies and calculations:

a) Anti-mouse FITC

- i) Horse anti-mouse FITC (Vector) 1:100 dilution
- ii) Alexa Fluor 488 Goat anti-mouse (Molecular Probes) 1:1000 dilution.

b) Anti-rabbit FITC

- i) Donkey anti-rabbit FITC (Jackson) 1:200 dilution
- ii) Donkey anti-rat FITC (Jackson) 1:200 dilution

c) Anti-mouse Cy3

- i) Goat anti-mouse Cy3 (Jackson) 1:200 dilution

d) Anti-rabbit Cy3

- i) Donkey anti-rabbit Cy3 (Jackson) 1:200 dilution
- ii) Donkey anti-rat Cy3 (Jackson) 1:200 dilution

Aspirate slowly to mix. Place on ice till application.

14. Remove (1° antibody) solutions from wells. Add 1mL PBS to each well and incubate 20 minutes at RT. Repeat once.

15. Add 200 μ L of 2° antibody solution (1.5 TBST, 5% NDS/NGS). Record wells and controls. Incubate 1 hour at RT.

16. Wash PBS, 1mL/well for 20 minutes. Repeat.

17. Remove plastic well using the provided tool. Continued...

Table 8.6: Immunohistochemistry – Cultured Cells continued.

Careful not to break slide or disturb cells.

18. Add antifade medium (Table 8.9)

19. Carefully place coverslip at an angle to the slide to avoid bubbles. Slide is now ready for viewing. Store at -15°C (dark) when not in use.

Table 8.7: Immunohistochemistry - Tissue

Day 1

1. Air-dry tissue slides under electric fan for 15-20 minutes. Label slide numbers with felt marker and record.
2. Using Coplin jars, soak slides in 1.5 TBST for 20 minutes.
3. Calculate antibody dilution according to data sheets. Add appropriate amount of blocking serum to 1.5 TBST buffer. E.g. 5% NDS or NGS: e.g. 40 μL /800 μL 1.5 TBST or 10 μL /200 μL 1.5 TBST (double amount for 10% blocking serum).
4. Add primary antibody. Mix slowly. Determine amounts using the following example: 2 $\mu\text{g}/\text{mL}$ dilution is need from an antibody solution with 0.136mg/mL concentration.
 - i) 200 $\mu\text{L}/\text{slide}$ is generally required.
 - ii) calculate ratio $(0.136 \text{ mg/mL} \times 1000 \mu\text{L/mL}) / 2 \mu\text{g} = 68$
 - iii) $200 \mu\text{L}/\text{slide} / 68 = 2.94 \mu\text{L}$ of antibody
 - iv) $200 \mu\text{L}/\text{slide} \times 5\% = 10\mu\text{L}$ of normal serum
 - v) $200 \mu\text{L}/\text{slide} - 2.94 \mu\text{L} - 10 \mu\text{L} = 187.06 \mu\text{L}$ of 1.5 TBST
5. Remove slides from Coplin jars. Dab and wipe back surface with Kleenex.
6. Wipe off buffer surrounding tissue with a Kimwipe.
7. Create barrier using hydrophobic marker (PapPen). Ensure thick line is created to prevent solution leakage during incubation.
8. Add 1 $^\circ$ antibody solution (1.5 TBST, 5% NDS/NGS). Do not dispense bubbles onto tissues as it will interfere with labeling within these regions.
9. Place in covered chamber and incubate at 4 $^\circ\text{C}$ overnight. Continued...

Day 2

10. Secondary antibodies and calculations based on 200 $\mu\text{L}/\text{slide}$ and add appropriate amount of 1.5 TBST buffer (5% NDS/NGS) with the following amount of respective secondary antibody:

a) Anti-mouse FITC

- i) Horse anti-mouse FITC (Vector) 1:100 dilution
- ii) Alexa Fluor 488 Goat anti-mouse (Molecular Probes) 1:1000 dilution.

b) Anti-rabbit FITC

- i) Donkey anti-rabbit FITC (Jackson) 1:200 dilution
- ii) Donkey anti-rat FITC (Jackson) 1:200 dilution

c) Anti-mouse Cy3

- i) Goat anti-mouse Cy3 (Jackson) 1:200 dilution

d) Anti-rabbit Cy3

- i) Donkey anti-rabbit Cy3 (Jackson) 1:200 dilution
- ii) Donkey anti-rat Cy3 (Jackson) 1:200 dilution

Aspirate slowly to mix. Place on ice till application.

11. Remove slides from covered chamber. Shake off excess solutions and place in Coplin jar containing 1.5 TBST buffer.
12. Repeat wash steps 3x, each with fresh 1.5 TBST.
13. Prepare 2° antibody solution outlined in step 10. Add drops carefully 200µL/slide. Incubate at RT for 1.5 hours.
14. Shake off excess and wash in 1.5 TBST 1x for 20 minutes.
15. Wash in 50 mM Tris-HCL solution x1
16. Remove slides from Coplin jars. Dab and wipe back surface with Kleenex.
17. Wipe off buffer surrounding tissue with a Kimwipe. Continued...
18. Add antifade medium, coverslip and place on cardboard trays. Slide is now ready for viewing. Store at 15°C (dark).

Table 8.8: SDS-PAGE Stock solutions

	<u>100 mL</u>	<u>250 mL</u>
<u>Acrylamide + Bisacrylamide (30%)</u> - foil and store at 4°C		
Acrylamide (Fisher, BP170-500)	30 g	75 g
Bisacrylamide (Sigma, M-2022)	0.82 g	2.05 g
	Dilute with ddH ₂ O	Dilute with ddH ₂ O
<u>10% SDS</u>	<u>100 mL</u>	<u>250 mL</u>
SDS (Fisher, BP166-500)	10 g	25 g
	Dilute with ddH ₂ O	Dilute with ddH ₂ O
<u>10% Ammonium Persulfate</u>	E.g. 30 µg + ~300µL	Made fresh daily
AP (BioRad, 161-0700)	Dilute with ddH ₂ O	
<u>1.5 M Tris (pH 8.8)</u>	<u>100 mL</u>	<u>500 mL</u>
Tris (Sigma, T-6066)	18.15 g	91 g
	Adjust to pH 8.8 with HCl	Adjust to pH 8.8 with HCl
	Dilute with ddH ₂ O	Dilute with ddH ₂ O
<u>0.5 M Tris (pH 6.8)</u>	<u>100 mL</u>	<u>500 mL</u>
Tris (Sigma, T-6066)	6.05 g	30.25 g
	Adjust to pH 6.8 with HCl	Adjust to pH 6.8 with HCl
	Dilute with ddH ₂ O	Dilute with ddH ₂ O
<u>Runner Buffer 5x (pH 8.3)</u>	1L	2L
Tris (Sigma, T-6066)	15.1 g	30.2 g
Glycine (Fisher, BP-381-1)	72 g	144 g
SDS (Fisher, BP166-500)	5 g	10 g
	Dilute with ddH ₂ O to 1x before use	Dilute with ddH ₂ O to 1x before use

Loading buffer (LB) 2x

	2.5 mL
0.5 M Tris (pH 6.8)	2.0 mL
Glycerol (Sigma, Y-3148)	2 mg
Bromophenol Blue (BioRad, 161-0404)	4.1 mL
10% SDS	(add before use to
β -mercaptoethanol (Fisher, G33-4)	10%)
ddH ₂ O	4.5

**Low MW prestained
standards**

(BioRad, 161-0305) Dilute to 1x with LB

**High MW prestained
standards**

(BioRad, 161-0309) Dilute to 1x with LB

Table 8.9:
Immunohistochemistry
Stock Solutions

<u>1.5 TBST</u>	<u>1 L</u>	
Trizma Base	6.05 g	
NaCl	15 g	
ddH ₂ O	950 mL	Make fresh use within 24 hours
Triton X-100 (Sigma, X-100)	3 mL	
	pH 7.6, make up to 1000mL ddH ₂ O	
<u>Tris (1M) stock</u>	<u>100 mL</u>	<u>500 mL</u>
Tris	12.10 g	60.5 g
	Dilute with ddH ₂ O	Dilute with ddH ₂ O
<u>50 mM Tris</u>	<u>500 mL</u>	
Tris (1M) stock	25 mL	
	Dilute to 500 mL	
<u>Anti-fade Medium</u>		
Glycerol	75 mL	
Parapheylenediamine	100 mg	
Tris (1M) stock	10 mL	
	Cover with foil and stir overnight	
	pH to 7.4, store at -20°C dark	

Table 8.10: Western-blot Stock Solutions

	<u>1L</u>	<u>2L</u>
<u>TBS</u> (pH 7.4, 20 mM Tris, 150 mM NaCl)		
Tris (Sigma, T-6066)	2.42 g	4.84 g
NaCl (Fisher, S271-3)	8.77 g	17.54 g
	Dilute with ddH2O	Dilute with ddH2O
<u>TBST</u> (pH 7.4, 20 mM Tris, 150 mM NaCl, 0.2% Tween20)	1L TBS	2LTBS
Tween20 (Sigma, P-5927)	+ 2 mL Tween20	+ 4 mL Tween20
<u>Blocking Buffer</u> (5% skim milk)	<u>50 mL</u>	<u>100 mL</u>
Skim milk powder (grocery)	2.5 g	5 g
TBST	47.5 mL	95 mL
<u>Blocking Buffer</u> (1% skim milk)	<u>50 mL</u>	
	10 mL 5% Blocking Buffer	
	+ 40 mL TBST	
<u>Transfer Buffer</u>	<u>1L</u>	<u>4 L</u>
Glycine (Fisher, BP-381-1)	14.4 g	57.6 g
Tris (T-6066)	3.03 g	12.12 g
SDS (BP-166-500)	0.5 g	2.0 g
ddH2O	Dilute to 800 mL	Dilute to 3200 mL
MeOH	+ 200 mL	+ 800 mL
	Do Not pH	Do Not pH
<u>Stripping Buffer</u>	2% SDS 100mM β -mercaptoethanol 62.5 mM Tris-HCl pH 6.7	<u>Submerge membrane</u> <u>Heat to 60°C for 30</u> <u>minutes</u>